JOURNAL OF CONSULTING PSYCHOLOGY

VOL. V NO. 4
JULY–AUG.
1 9 4 1

CONTENTS

HAROLD E. BURTT — FRONTISPIECE

MARKET PROBLEMS AND MARKET RESEARCH

CURRENT TRENDS IN MARKETING RESEARCH — HAROLD E. BURTT — 141
PROVING GROUNDS ON PUBLIC OPINION — HENRY C. WEAVER — 149
PROBLEMS OF SAMPLING IN MARKET RESEARCH — FRANK STANTON — 154
CHARACTERISTICS OF THE QUESTION AS DETERMINANTS OF DEPENDABILITY — JOHN G. JENKINS — 164
EVALUATING THE EFFECTIVENESS OF ADVERTISING BY DIRECT INTERVIEWS — PAUL F. LAZARSFELD — 170
EFFECTS OF REPEATED INTERVIEWING ON THE RESPONDENT'S ANSWERS — FLOYD F. RUCH — 179
THE MUSEUM TECHNIQUE APPLIED TO MARKET RESEARCH — GEORGE A. BENNETT — 183
THE ROLE OF PSYCHOLOGICAL INTERPRETATION — ARTHUR W. KORNHAUSER — 187

NEWS NOTES — 190

PROGRAM OF THE FIFTH ANNUAL CONFERENCE OF AMERICAN ASSOCIATION FOR APPLIED PSYCHOLOGY, NORTHWESTERN UNIVERSITY, SEPTEMBER 1–3, 1941 — 191

AMERICAN ASSOCIATION FOR APPLIED PSYCHOLOGY

VOLUME V JULY-AUGUST, 1941 NUMBER 4

JOURNAL OF CONSULTING PSYCHOLOGY

PUBLISHED BY THE AMERICAN ASSOCIATION FOR APPLIED PSYCHOLOGY, INC.

EDITOR

MRS. J. P. SYMONDS, 525 West 120 Street, New York, N. Y.

ASSOCIATE EDITORS

HAROLD E. BURTT, PH.D., The Ohio State University
HAROLD E. JONES, PH.D., University of California, Berkeley
ARTHUR W. KORNHAUSER, PH.D., University of Chicago
BERTHA M. LUCKEY, PH.D., Board of Education, Cleveland, O.
CARL R. ROGERS, PH.D., The Ohio State University
GOODWIN WATSON, PH.D., Teachers College, Columbia University
F. L. WELLS, PH.D., Department of Hygiene, Harvard University

The JOURNAL OF CONSULTING PSYCHOLOGY is the official journal of the American Association for Applied Psychology, Inc., which was organized in 1937 as an amalgamation of the Association of Consulting Psychologists, Inc., and the Clinical Section, American Psychological Association, Inc., with the affiliation of regional associations for applied psychology.

The JOURNAL OF CONSULTING PSYCHOLOGY publishes results of research bearing on the application of psychology, short articles and communications relating to professional psychological service, official reports of the activities of the American Association for Applied Psychology, Incorporated, and of other organized groups of professional psychologists with representation in its Board of Affiliates, information and notices of interest to psychologists, resources available to the profession, and critical reviews of the professional literature. Contributions are invited from all fields which are significantly related to applied psychology.

EDITORIAL OFFICE

Manuscripts, general correspondence and publications intended for notice or review should be sent to Mrs. J. P. Symonds, *Editor*, 525 West 120th Street, New York, N.Y. Manuscripts are accepted only upon recommendation of the editorial staff. Proof will be submitted to the author at the time of publication. Corrections made by the author are chargeable to him. Manuscripts should be submitted in duplicate. References to publications should conform to the style of this JOURNAL, giving in order: name of author, title of article, full name of periodical, year, volume, number, pages. References to books should give in order: author's name, title of book, place of publication, name of publisher, year, pages. Articles in this JOURNAL are indexed in *Psychological Abstracts*.

BUSINESS OFFICE

J. W. DUNLAP, PH.D., *Business Manager*, University of Rochester, Rochester, N.Y.

Correspondence regarding subscriptions and business matters should be addressed to the publishing office at Colorado Springs, Colorado, or to Dr. Jack W. Dunlap, *Business Manager*, University of Rochester, Rochester, N.Y.

The JOURNAL OF CONSULTING PSYCHOLOGY is published bi-monthly, six times a year. All members of the American Association for Applied Psychology are entitled to receive the JOURNAL upon payment of their annual dues of which two dollars is applied to the Journal. Affiliated state or regional Associations may arrange a group subscription with the Business Office through their Representative to the Board of Affiliates whereby members of the group subscription are entitled to receive the JOURNAL upon the payment of two dollars added to their Association's dues. The individual subscription is $3.00 a year and single copies are 60 cents a number. Postage is prepaid. Claims for non-receipt of a number will be honored in the Business Office within two weeks after receipt of the next succeeding number.

Copyright 1939 by American Association for Applied Psychology, Inc. Printed at 19 West Colorado Avenue, Colorado Springs, Colo., and entered as second-class matter at the Colorado Springs, Colo., post office.

Harold E Burtt

THIS IS ONE OF A SERIES OF PORTRAITS OF PSYCHOLOGISTS WHO HAVE CONTRIBUTED NOTABLY TO PROFESSIONAL PSYCHOLOGY

Volume 5 JULY-AUGUST, 1941 Number 4

JOURNAL OF CONSULTING PSYCHOLOGY

CURRENT TRENDS IN MARKETING RESEARCH

By HAROLD E. BURTT, PH.D.
OHIO STATE UNIVERSITY

UNTIL recent years marketing was motivated by a one-sided philosophy. Its objective was to sell the product regardless of the consumer's interest or preference. In personal selling this was typified by the drummer of a generation ago with his breezy manner, his glib conversation and his stories. He could sell *anything* and did it largely by sheer force of personality. In advertising the tendency was to pass the problem along to the advertising agency. They suggested, "Let Us Sell Your Product," regardless of whether the public wanted it. Thus the policy developed of manufacturing anything that occurred to the promoter and then marketing it in the face of opposition, breaking down sales resistance, applying high pressure and employing sure-fire appeals.

With increasing competition it became evident that this procedure was not adequate. Manufacturers began to clutch at straws which would enable them to survive the competition and they began to call upon science for help. Interest began to be focused upon the consumer. Some of this interest dealt with economic factors such as buying power, income, and population distribution. But Psychology also began to enter the picture with a consideration of the consumer's desires and preferences. It was proposed to adjust the program somewhat in the light of these desires in the hope of securing larger sales, increased satisfaction with the product and resultant good will for the manufacturer. This kind of program implies a research approach because it is necessary to secure some notion of the direction of the consumer's desires and this cannot be done by intuition but must be approached empirically. Programs of this sort have been directed at numerous problems. For instance interviews are conducted to ascertain preferences for the different brands of a product in order to see what advertising has done to these preferences. Similarly studies are made of radio listening practice,—programs preferred and also programs tuned in with the latter aspect as a check on the former. A concern is sometimes interested in the attitude of the public toward it and gets the answer by concealing the name of the concern in a more extensive attitude scale. Before a new package is adopted for a product consumers may be interviewed as to the shape, size, and decoration that they prefer. Interest may center on buying

habits in a given field, for instance as to whether one buys motor oil at the same place he buys his gasoline. The investigation may even center upon rather minor differences in the product as affecting the consumer's preference, as for example shape of fenders or radiator on a car.

Some commercial organizations have been set up on a nationwide basis to conduct research studies like the foregoing for their clients. The principal one directed entirely by psychologists is the Psychological Corporation. Some industries have organized research departments of their own to deal with such problems.

The methods employed in these research programs are mainly interviews and questionnaires. The latter constitute an old story for the psychologist but recent years have brought out important technical improvements in questionnaire construction. Interviewing projects are playing a large role in our culture. The initial impetus came from interest in magazine and newspaper reading habits but now the technique has been extended to the broad field of public opinion. There is still, however, a great deal of work being done on market problems by these procedures. Psychology has some contribution to make in the selection and training of interviewers, in the formulation of questions and in determining whom to interview.

Thus a survey of some of the psychological aspects of market research is timely and the present number of the JOURNAL includes a discussion of these by persons who have been active in this field in recent years. It is appropriate to hear from Weaver representing one of the large industrial concerns which has pioneered in this type of work and has been conducting its own research program for years. Doubtless many of the readers of this JOURNAL have received the "Proving Ground of Public Opinion" circulated by General Motors. Much of the work of this organization has centered around these mailed questionnaires which have included some original features such as the attempt to increase returns by livening the blank with thumb-nail sketches and conversational style. It may be noted that their sampling problem is somewhat different from most other projects because they want only automobile owners and these data are available from registration lists. Weaver indicates also that his organization on occasion uses interviews and likewise organizes smaller groups of respondents who will answer questions repeatedly,—the Panel technique which is mentioned in some of the other articles in this series.

We find three of the articles dealing with technical problems in market research. The first of these by Stanton involves the ubiquitous problem of sampling with especial reference to the representativeness and the size of the sample. Some commercial workers in this field are unaware of the importance of the former and of the statistical possibilities in interpreting differences as to their significance. Increasing the number of cases will not necessarily reduce the sampling bias. The controlled or stratified sample is generally the best. In case of doubt as to whether a given variable might affect the result it is well to take account of it in drawing the original sample. It is possible to make a further check on the sample by including in the interview questions regarding the variables used in selecting the sample. Careful attention to the two main points of this article would go a long way toward increasing the reliability of many market research studies.

Continuing technical considerations,

Jenkins contributes a much needed discussion of characteristics of the questions themselves. Much loose work has been done in constructing questionnaires and schedules on the basis of personal opinion or hunches. The necessity of pretesting the questions is very wisely emphasized. Among the difficulties encountered the following should be noted: (1) biasing the answer by the wording of the question or by restricting the number of alternative answers; (2) failing to direct the response sufficiently, that is, making a question indefinite or ambiguous; (3) exceeding the respondent's ability by using large words or complex questions; (4) inviting inaccurate responses by questions which might be embarrassing or incriminating. Psychologists have contributed notably by calling attention to considerations such as the foregoing in the actual construction of the questions used in market research projects.

Lazarsfeld devotes himself to a further technical consideration of the type of survey that is concerned with *why* the respondents do some particular thing, for example make a purchase. His article centers largely around a typical case of this sort where the problem was to determine if advertising was the cause of the purchase. He suggests, as do some of the other writers in this series, a pretest of questions with a view to determining various alternative reasons for the purchase. These alternatives may be embodied in check lists to be employed in the final interview. He proposes to test the validity of the reasons by further questions, for instance, if price is mentioned, to determine whether the persons were exposed to advertising regarding price. He shows how interviews may be made profitably in two trade areas one without advertising and one with advertising for a given product.

Looking at the difficulties which may arise in the above procedures, Ruch raises a question which has long bothered some of us, namely the effect of repeated interviews on the same individual. He suggests that either the respondent may become hostile or may actually become interested and have better rapport. He approaches the problem largely by analogy noting the effect of familiarity on preferences, indications of how memory for one interview influences a subsequent one, and how motives may effect preferences as shown by efforts to mark a personality schedule the way a particular vocational group would mark it. His discussion leads up to making a case for the Panel technique in which persons who have expressed sufficient interest may be interviewed repeatedly. Further study of the problems raised would certainly be helpful.

One article in this series, that by Bennett, opens up a quite different problem. It deals not with consumers' preferences and techniques for ascertaining them but rather with commercial displays and the problem of how to direct a visitor to certain portions. He takes his cue from studies of behavior in a museum where the visitors are observed and their behavior recorded. Tendencies are noted, for example, to follow the right wall or to devote attention intermittently, lingering for some time at one point and then jumping over several items to the next source of attention. He shows how this technique has been used in some commercial exhibits and suggests some further possibilities.

Finally, we have a note of caution by Kornhauser on the matter of interpretation. He recommends such interpretation not merely ex post facto but at

the outset in setting up the problem. He urges a wholesome combination of theory and practice. The former may suggest interpretations which can be checked empirically in individual cases. For example, some theoretical notions regarding preferences as related to age may be checked with reference to a specific product under investigation. Theoretical lists of motives even though not agreed upon universally may be useful in suggesting facts to seek in an objective study. Principles may be drawn from other fields, for example, psychoanalysis, which actually may play a role in attitudes toward comic strips or moving pictures. He wisely suggests consideration of the entire behavior area and various alternative explanations so that none may be overlooked.

The articles in this series thus show that the psychologists are giving some serious thought to market research problems. We have passed the initial stage of hastily constructed questions, we are pretesting our techniques, selecting our samples carefully, attempting to take into account all the possibly significant variables and exercising care in the interpretation of results.

PROVING GROUND ON PUBLIC OPINION

By HENRY G. WEAVER

DIRECTOR CUSTOMER RESEARCH STAFF

GENERAL MOTORS CORPORATION, DETROIT, MICHIGAN

KENNETH GOODE once laid down this very simple formula for business success:

"Find out what people LIKE and do MORE of it."

"Find out what people DON'T like and do Less of it."

In a small business, or what we might call a "one man shop," where the proprietor is in intimate daily contact with the public whom he serves, the likes and dislikes of the customer are learned without conscious effort, but when a business attains large proportions there is created a gap between those responsible for directing the affairs of the business and the ultimate consumer upon whose patronage the business must depend. And yet the need for understanding the likes and dislikes of the customer are just as great in a big business as in the "one man business." If anything, the need is even greater for two reasons:

1st. People appear to be somewhat more exacting in their dealings with a large corporation than they are in dealing with a small business.

2nd. A large business, if it is to take advantage of the economies resulting from large scale operations, must lay its plans and make its commitments on a long range basis without much opportunity for day to day revision.

Robert F. Elder[1] has amplified this latter point with particular reference to the manufacturing business, as follows:

> Under a system of manufacturing goods to the consumer's order, the problem of proper design assumes small importance to the producer. With the increasing tendency to manufacture for stock, the producer's design problem becomes acute. In the former case the consumer exerts a positive and active influence; in the latter case his influence is purely negative and passive. The buyer of ready-made goods is seldom articulate in his criticisms. He expresses his judgment by purchasing or by refusing to purchase. If he does become vocal with complaints or suggestions, the length of the distribution process, with its numerous intermediaries, usually prevents their getting to the producer's ears.

As regards this latter point, I would like to advance the thought that the length of the distribution process is not the only cause. An equally important point, in my opinion, is that the distribution process, or rather the sales process, is somewhat in the nature of a "one way circuit." A good salesman, either wholesale or retail, is not likely to be effective as a medium for relaying criticisms. The very fact that he is a good salesman is likely to make him ineffective when it comes to reversing his role and reporting the negative aspects.

The foregoing are some of the considerations that have led to a fairly general trend in large scale operations to set up independent means for appraising customer likes and dislikes. Since General

[1] Product Design for the Market. *Mechanical Engineering*, August, 1932, p. 543.

Motors is a pioneer in this field and since my experience has been confined almost entirely to General Motors problems, I think that the most useful thing I can do in this article is to review certain aspects of our own work.

CUSTOMER RESEARCH ON PRODUCT DESIGN

The General Motors Customer Research Staff was set up as a self-contained department in 1933. In the course of a year we make from twenty to thirty surveys, contacting with over two million motorists, owners of all makes of cars, located in all sections of the country. These surveys relate not only to features of mechanical design, styling, etc., but to various aspects of advertising, sales promotion, service, etc. Right here I'd like to emphasize the fact that, contrary to the general impression, we employ the personal interview technique as well as direct mail. As I see it, the two techniques are complementary to one another and they both have their place in any well-balanced customer research program. It is true that we, in General Motors, depend somewhat more on direct mail than is the general practice. The reasons are as follows:

1. It is necessary to do our sampling not just among rank and file automobile users — but on a carefully controlled basis of selection as regards make of car owned, year model, body style, territorial location, etc.
2. It is very expensive to get an adequate and properly balanced sample through personal interviews whereas such data are readily available through mailing lists built up from automobile license registration data.
3. The buyers of new motor cars are found mostly in the higher economic strata. Generally speaking, they are people to whom writing is not much of a burden.[2]
4. With properly developed techniques we are able to get high returns through the mails.

Our surveys relating to the product itself disclose the *relative popularity* of different features and characteristics of design, and by comparing the results year after year, we are able to measure the *trend* of public preference as regards each item—whether it is becoming *more* popular or *less* popular. Furthermore, since our questionnaires are mailed in large quantities, it is possible to study variations in preferences as between different sections of the country and different types of communities—as well as by price classes. In addition to the questions on specific features of design, the questionnaires used in these general surveys invite the respondent to "cast his vote" on broad overall questions of operating characteristics or what we might call the *results* or the *effects* of design. In other words, the motorist is invited to "rate" the degree of importance that he attaches to such things as: dependability, economy, comfort, safety, ease of control, performance, appearance, etc.

The data resulting from this type of question, measured against the results of General Motors Proving Ground and Laboratory tests, indicate the degree to which the characteristics of current models are in line with the trend of public desires. They also serve as "straws in the wind" suggesting the future course of progress in a general way, although considerable caution must be

[2] Contrary to the general impression, there are only about seven or eight million new car buyers in the United States. The balance of the cars are owned by people who bought them as used cars. In our researches among used car buyers, we are relying more and more on personal interviews.

exercised in interpreting such data into terms of engineering design. The expressed tastes and desires of the consuming public must always be weighed against the technical aspects and the considerations of cost in relation to the advantages to be gained. Changes in one direction may frequently necessitate changes in other directions. The problem is far from simple. The development of any *one* characteristic or quality to an extreme degree, is usually attended by sacrifices in some other direction and it is not to be expected that the layman could anticipate the net overall result. In other words, the maintenance of a consumer research activity does not mean a blind acquiescence to the expressed tastes and desires of the consumer, which, as applied to a highly technical product, such as the motor car, might lead to violation of sound engineering, to constructional difficulties, or to increases in cost outweighing the advantages to be gained. After all, the consumer is more of an expert on use than he is on design. Therefore, he can hardly be expected to project himself very far beyond that which he sees and feels in his day to day experiences.

In recognition of this fact, we have, over a period of years, as an incident to our general surveys, built up a panel of "special correspondents" or what we call "Motor Enthusiasts" restricted to the names of people who make motoring a hobby—who take more than a passing interest in problems of engineering design—and who *do* have the ability to project their thinking somewhat beyond what they actually see in the cars of today. Since these "Motor Enthusiasts" think in advance of the average car owner, their reactions are more significant as bearing on the future. The questionnaires returned by the Motor Enthusiast group are compiled separately in order that the results may be studied in comparison with the voting of the mass group.

Along this same general line of thought, I'd like to say that Customer Research, if it is to provide dependable guidance for future planning, must aim to get two types of reactions.

1. *Offhand Reactions* — what is the existing state of the public mind as regards a given proposition?
2. *Deliberative Reactions* — what are the reactions to the same proposition after the pros and cons are brought to the attention of the respondent?

The former aims to take a cross section of public opinion just as it exists—without reference to the soundness or unsoundness thereof. The latter type concerns itself not with *existing attitudes* but with what the attitudes will most likely be *after* the motoring public — through experience or otherwise — has come into possession of the full and complete facts. While there can be no question but that spontaneous reactions are highly significant as bearing on the short haul, deliberative or well thought out reactions are likely to be more significant as bearing on the long haul or ultimate trend of public thinking.

In line with this theory we supplement our general questionnaires with a series of special questionnaires, each dealing with a separate feature of design—and we discuss it in considerable detail and present the pros and cons before calling for a vote. Such auxiliary inquiries, which are directed in the main to the so-called Motor Enthusiast group, help to "sharpen the sights" as regards the most logical trends for the future. They also throw valuable light on the "boundary lines of public acceptance" as regards specific features of de-

sign—providing the answers, or at least the partial answers to the question: Which of the projects that we have "on the fire" in our Research Laboratories and Division Engineering Departments, which of them would be most logical to push forward into production? I might say in passing that these special questionnaires, seeking deliberative reactions, are much more difficult to prepare than a questionnaire which merely mentions a feature and calls for a vote. Although, generally speaking, our mailings on the deliberative type of questionnaire are quite small, the returns run high and the voluntary comments are especially helpful because they provide the atmosphere back of the attitude.

CUSTOMER RESEARCH AND MERCHANDISING

The foregoing relates to consumer researches on questions of product design. We will now touch upon the merchandising aspects. In the final analysis, a good product, or a good feature of design, supported by good service will, within itself, win public acceptance — that is over a long period of time—but from a standpoint of practical merchandising in a fast moving, highly competitive industry we can't afford to wait for nature to take its course. The big job in merchandising is to speed up the process of public acceptance which, if we do merely sit by and wait for nature to take its course, may require a period of time approaching the life of the product—and that, of course, is too long. Assuming that the product has been designed sensitively in line with customers' needs and desires (consistent, of course, with sound engineering), then the next question is : How are we going to establish the facts of its goodness in the mind of the consuming public?

Before we can set about to do an intelligent job of influencing the thinking of the consumer, we've got to know what he *already* thinks and why he happens to think the way he does think. To a considerable degree such information is provided as a "by-product" of our surveys on constructional features, because most of the questionnaires used in such surveys are designed to dig beneath the surface of the consumer's attitudes and throw light on the whys and wherefores *back* of his likes and dislikes.

To supplement such information we make special surveys designed to throw factual light on the basic underlying problems of consumer psychology, aiming to provide fundamental information upon which to build greater efficiencies into the merchandising procedures—just as the engineering handbooks provide basic information which serves as a foundation for mechanical design: What are the attitudes and attributes of the consumer as affecting his habits of purchase and consumption?

Why do people buy automobiles? What are the most potent influences that bring about the decision to buy? How do they set about the job of choosing? How far ahead of time does the typical customer make up his mind to buy? What are the forces at work during the period of "incubation"? Which of these forces are the most important?

What do the answers to such questions indicate as bearing on the broad problems of merchandising and customer relations? What's to be done about it? How can it best be done? By *whom* and through what *tools?*

Right here it is well to point out that Customer Research is not a specialized activity—that is, not in the usual functional sense. Any attempt to restrict its operations to anything less than *all* those phases of a business which affect

the customer—either directly or indirectly— would be contradictory to the broad objectives of the work.

The purpose of Customer Research is to give a true and complete reflection of consumer thinking. Consumer thinking does not follow the pattern of a functional organization chart. The consumer thinks in terms of net results— i. e., how he himself, as an individual, is affected by the product and the policies. Therefore, Customer Research, if it is to be fully effective, must approach its job from the same viewpoint—namely, a viewpoint that is from the *outside looking in*—observing and recording the composite or net effects of ALL specialized activities on the attitudes and actions of the buyer. In other words, Customer Research, in its contacts with the consuming public, must lend an attentive ear to *any* and *all* kinds of reactions without reference to the more or less arbitrary boundary lines of departmental organization. Then as a matter of routine procedure and working from *effects* back to *causes*, Customer Research should attempt to segregate and classify the reactions and pass its conclusions to the proper departments.

Because of its non-specialized and unrestricted approach, Customer Research has the opportunity to serve as a coordinating influence — somewhat in the nature of an antidote to the specialized viewpoint—and in the process it brings to light opportunities for new improvements and new developments in the "no man's land" lying *between* the highly specialized departmental functions.

PROBLEMS OF SAMPLING IN MARKET RESEARCH

By FRANK STANTON, PH.D.

DIRECTOR OF RESEARCH, COLUMBIA BROADCASTING SYSTEM

IN the early thirties when market research was first becoming popular with advertisers and advertising agencies, one commercial research specialist, when approached by a prospective client with an inquiry as to the size of the sample necessary for a specific study, would reply invariably, "How much money do you have to spend?" and the sample was prescribed according to the answer!

That was ten years ago. There are some in the field who still operate on such terms but, for the most part, market research technicians have profited by the past ten years' experience. Today, while sampling is still one of the major problems in research, we are much more skilled in dealing with it (13, 9, 15).[1]

SAMPLING PROBLEMS

Two important questions face every worker in this field: *Is the sample representative?* and, *Is the sample large enough?* To answer these two questions satisfactorily requires previous knowledge of the universe under study (item one: *representativeness*) and the application of statistics (item two: *size*). Far too few of today's market research studies meet even the minimum standards of *representativeness* and *size*. In fact most studies make little or no mention of self-imposed checks on the representativeness of the sample. Neither do they indicate, even with the cruder statistical measurements at their disposal, anything regarding the sample's adequacy in terms of size.

There are several probable reasons for this apparent laxness with regard to sampling problems. In the first place, as in other fields of research, good work costs money. And business today has not yet been educated to discriminate between good and bad market research; nor has it, therefore, come to appreciate costs. But to insure effective sampling requires time and facilities. In many cases where insufficient time is allowed for careful sampling, it is not necessarily a matter of costs but stems usually from the very nature of the situation involved. For example, a new product is launched and an immediate check is required to determine the advertising's effectiveness. In such a problem the effects are not static; they must be ascertained immediately, in many cases without benefit of experimentation and "pretesting."

Another probable reason for present-day sampling weaknesses is the complete lack of understanding or interest on the part of most business executives for the problems involved. Without a knowledge of the theory of sampling and the consequent "sampling fluctuations," they cannot be expected to appreciate the fine points of sampling and are, therefore, unwilling to wait and pay for adequate checks. There are exceptions, of course, to this charge. The *Literary Digest* fiasco opened the eyes

[1] For detailed discussions in this field, any one of these recent articles is recommended.

of many. Similar dramatic reversals within specific industries have either forced research out entirely or improved it. The difficulty is that in many instances there are insufficient methods of validation and, therefore, little opportunity to check results except by statistical devices and "post interview" controls of the sample, neither of which is entirely foolproof. Market research has few natural opportunities to validate the accuracy of its findings. Unlike public opinion research it does not have the benefit of national or state and local elections for checking its workmanship.

Probably the basic reason for the lack of sampling safeguards lies within the research personnel itself. A large number of independent research organizations are operated by business men and *not* by research men. In other words, the men at the top are not trained research-wise. Neither through professional schooling nor experience have they mastered the techniques of sampling and all it implies. This indictment should disappear as fast as universities take over the field on an effectual basis. Also it will tend to correct itself as research is taken more seriously by business. Once it gets beyond the sales promotion and advertising phase and is used intensively by and for management, the incompetent technicians will be thinned out and a new generation will emerge.

Is the sample representative?

Generally speaking, in terms of use, there are three types of sampling techniques: *accidental, random* and *stratified* or *controlled.* In discussing the representativeness of a sample, each of these three conditions will be reviewed briefly.

The usual market research study can be classified in the group depending upon accidental sampling procedures. For sales purposes, research organizations usually describe the sampling as random, but, as we shall see later, the usual consumer survey violates the definition of random sampling at every turn of the way. Among the excuses cited by the practitioners of this accidental method would probably be the factors of cost and ease. While the accidental "plan" of sampling has nothing whatever to recommend it as far as methodology goes, it should be noted here that there are wide variations in its application. In some studies, accidental sampling may mean interviewing anybody; in others, investigators may attempt to institute some controls by trying to interview what they consider a "cross section."

SAMPLING PROCEDURES

In the usual market research study, insofar as sampling is concerned, three steps are involved. First, the market or markets to be studied are selected. Second, the universe within these markets is determined according to the subject of the study. If the problem concerns automobile tires, then the universe is confined to car owners. On the other hand, if the study involved the use of cosmetics, the interviewing would have been confined to women. The third step in the accidental plan is to turn the interviewers loose with the questions and allow them to query any member of the universe with little or no instructions.

Obviously there are many objections to such a procedure. Without a plan for interviewing, based either on a "random list" or quota controls by "stratified groups" (both to be discussed later), the sampling left entirely to the investigator will be biased. It thus happens that the investigator who follows this method of sampling is frequently misled as to the nature of the sample. The effort of the investigator to exercise no con-

trol over the selection of the sample does not mean that no control was operative.

The tendency on the part of interviewers to contact only certain restricted areas or groups within the total statistical population results in a rather carefully selected sample. The term "accidental" thus becomes a misnomer because clear-cut influences are at work. These factors are such that an experienced research worker can predict the nature of the sample selected in this way with amazing accuracy. Usually what happens is that the interviewer works in his or her own age, sex, income group or section of the market. Under such conditions the reports will not represent all social groups unless the investigators themselves are selected from all strata.

ACCIDENTAL SAMPLING

In the accidental "plan," most of the interviewing is done where it is easiest and that usually is in the middle socio-economic range of society. Insufficient attention is given to the poor and relief families, negro and foreign born and also, although not to the same extent, high income groups are excluded. If the problem under study is related to social or economic factors, results based on such sampling are ineffective—or what is even worse, misleading.

It is difficult to estimate what share of today's market research is conducted along this line. In the early days, it represented most of what was done but since the studies then were confined to relatively few markets, there was superior supervision and hence probably better work. With the "advance" in market research, studies have spread geographically until today most organizations maintain field workers on a loosely organized nationwide basis. Because of this spread, there is less control and supervision; most of it is by mail in place of the personal and direct contact of old.

It should be pointed out that the applications of the *accidental* method have not been confined wholly to market research. Most sampling for research conducted by investigators associated with schools and colleges is carried out along accidental lines. (And since the ordinary student body is further from the mean of the total population on most factors, the results obtained are usually atypical of anything save the sample examined.) Here the problem results probably from lack of opportunity to extend the sample. The serious thing, however, is that uncritical writers accept the findings of such studies without qualification as to the universe observed. There is ample justification for the contention that the chief contribution of the research done in colleges and universities to the field of market research has been a demonstration of the futility of attempting to generalize regarding the total population from data which have been collected in experiments using only college students as "subjects."

RANDOM SAMPLING

In the application of the true random procedure, it is assumed that the universe to be studied is listed or catalogued. In such instances, if a 10 per cent sample is to be taken, the randomness of the sample is achieved by interviewing every tenth name in the catalogue. This should yield a virtually random sample provided the names have not been arranged according to any classification except an alphabetical one.

Surveys dependent on unlisted universes cannot follow this plan of operation. Or for that matter, it is impracticable, even where catalogues exist, for

personal interviews because it involves wide travel to locate individual cases. Where it is used at all in every day market research, it is usually confined to telephone or mail studies, and only then if the universe is completely listed or catalogued.

One application of the random technique is the radio program measurement plan known as the "coincidental telephone" method. Here the universe is composed of radio-telephone homes in selected cities. (This plan is the basis for the program rating service maintained by C. E. Hooper, Inc.) In the telephone "coincidental interview," information is obtained regarding radio-set use at the time of the call. The operator notes whether the set is reported in use and, if in use, to which station and (or) program the set is tuned. Additional data are obtained in special cases on audience composition by age and sex and data on "sponsor identification" of the program being received. So much for completed calls. Records are also kept of all other initiated calls according to "non-answer," "refused," "busy," "sets not in use," etc.

The sample is drawn by selecting every *nth* resident telephone family listed in the telephone directories for the cities studied. In this way, it is assumed that the random selection insures a representative sample of telephone-radio homes. Insofar as conclusions regarding listening are not projected beyond the universe sampled, this procedure is satisfactory.

One can readily guess what usually happens to data collected under such conditions: projections are made to larger (e.g., non-telephone) universes without adequate evidence that the telephone-radio homes surveyed are representative of *all* radio families. It is reasonable to assume, however, that such is not the case for all programs because the *appeal* of the program, the *hour* of the broadcast and *cities* surveyed all may influence the audience in non-telephone homes differently from the listeners in telephone homes.

Program appeal. For the most part, telephone ownership is confined to upper income groups (29,300,000 radio families; 12,500,000 telephone families). A program which appeals more to one-half of the income scale than the other would be incorrectly estimated if projections were made. Below are examples of actual radio program ratings in telephone and non-telephone homes. (Total radio families in each group equals 100 per cent.)

	Telephone	Non-Telephone
"Johnny Presents" . . .	20.6%	26.1%
"Second Husband" . . .	8.6	15.6
"Tune up Time"	14.5	12.8

Hour of broadcast. Again, because telephone ownership correlates positively with income and because the periods of listening will vary income group by income group—due to working and living habits—this factor, too, operates to make projections unsafe from telephone homes to the total radio audience.

Cities surveyed. The markets surveyed are selected because they all receive program service from three or more networks and have a sufficient number of telephone families. Therefore, projections to other geographic areas are unsound due to differences in program service (signal dissipation) and variations in program appeal by different population groups. Below are the same programs listed above rated by "city-size groups." (Total radio families in each group equals 100 per cent.)

	Cities over 100,000	Cities under 100,000	Rural
"Johnny Presents"	22.4%	26.5%	25.3%
"Second Husband"	10.9	10.7	9.8
"Tune Up Time"	13.9	11.7	8.0

Insofar as the telephone coincidental technique employs the random sampling procedure correctly, it yields an un-

biased sample. But data derived from a radio-telephone universe may not be projected to a larger universe unless it can be demonstrated that the smaller universe has listening habits identical with the larger one, i.e., the smaller population (radio-telephone families) is representative of the larger population (all radio families).

In connection with this discussion of random sampling, it should be pointed out — in contradistinction to some other writers — that a random sample is not necessarily representative. Results based on random samples tend to distribute themselves about the true measure in a normal distribution. As in any normal distribution, therefore, one should expect deviate cases. Just as in a game of poker, where random sampling is practiced, there are cases where one hand is composed of a "royal flush." Such a hand is random but not representative. It is the infrequent deviate case.

Mail studies are usually resorted to in the interests of cost or ease of administration. But unless the replies approach 100 per cent of the original random mailing, the results may not be considered representative. Sometimes it is necessary to use a mail questionnaire in market research to reach special groups scattered throughout the total population. For instance, in querying school teachers regarding classroom use of radio, a random mailing is more efficient than personal interviews provided a complete list of teachers is available. It is only safe from a sampling point of view, however, when an effort is made to hear from all persons "sampled." Or, putting it another way, it is unsafe to project to the total universe because there is little likelihood that those who reply are representative of those who do not reply.

In a mail study of classroom use of radio, an analysis was made of the percentage of radio ownership among respondents to the original mailing of the questionnaire and those who were sent a special follow-up form (12). The figures below show clearly how dangerous it is to project from the usual mail return.

	Original Mailing	Follow-Up Mailing
Teachers with classroom radios	33.0%	23.8%
Total teachers heard from, (100 per cent) . .	1,402	1,080

In this study there was a third group composed of those teachers who did not answer either mailing. While it was unsafe to project to the total sample, it was not unreasonable to assume that teachers from whom no response was received owned even fewer radios than the ones who replied.

In reality, while random sampling meets the theoretical standards of representative sampling, it is very difficult to apply the principles of the random method because seldom is the universe to be sampled catalogued, or when it is, the actual interviewing is impracticable on a face-to-face basis because of the travel involved. And if mail is used, it is unsafe unless all cases are heard from. The telephone coincidental interview used in rating radio programs in restricted areas is one of the few practical applications of the method.

CONTROLLED SAMPLING

The "controlled" or "stratified" sampling procedure represents the most desirable plan for insuring representativeness. In this method, the interviews are allocated by strata according to the proportion each stratum represents in relation to the total universe. If the problem under study is related to income—

and if 10 per cent of the universe is in the top income group, 25 per cent in the upper middle group, 30 per cent in the lower middle and 35 per cent in the lowest group—then the sample would be "controlled" or "stratified" to the same proportions. Of course, it does no good to stratify the sample to insure representativeness unless one knows by prior test that the variable under study is related to the particular plan of stratification involved. For example, if one were studying preferences for radio program types, he might want to "control" or "stratify" the sample by age or economic levels. That is because we know program types vary by age and economic status. On the other hand, no effort would be made to stratify the sample by height or weight groups because such factors probably bear no constant relationship to program types.

It is very dangerous to omit the control on certain variables when there is the slightest suspicion that they are related in any way to the problem under investigation. Marital status, for example, might, on the surface, appear to be unrelated to a certain market problem. It might be argued, however, that marital status is an important determiner of buying, listening or reading habits. In any event, the researcher would do well to control on this variable in selecting his sample. Of course, this principle can be carried to the point of impracticability, but an acknowledgment of its validity is a step in the right direction.

Under certain circumstances, the allocation of interviews or sampling is controlled by more than one variable. As an example, the problem of measuring a radio program's audience within a cluster of counties might be considered. Here at least two factors should be controlled. First, there is the question of stratification by income groups (*vertical* sampling control) to get the proper representation from each economic level. This is necessary because program appeal and listening behavior is related to economic level. Second, since the program's audience is to be measured over a group of counties, the proper stratification by population groups and geographic areas (*horizontal* sampling control) must be maintained. This is because a program's appeal varies by size of town and the opportunities for listening to it vary by location in the area because a radio station's signal is not constant in all places.

The plan with a two-way "control" or "stratification" is to decide on the proportion of interviews in each place (*horizontal* sampling), and then within each place—city, village or rural area—establish the quotas or proportions of interviews by income groups (*vertical* sampling). To control one of these and not the other may introduce sample bias into the results.

Once the quotas or allocations of interviews are set up for each stratum, the investigators are directed to obtain a cross section of each level (*vertical* sample), in each place (*horizontal* sample). Sampling within each stratum is as random as possible. Usually, however, in technical terms, it amounts to accidental sampling within each stratum. But this is much more representative because the final total sample will have the proper number of interviews from each group. By setting up the controls or stratification, the sample is forced into proper balance and is, therefore, more representative.

This plan of stratified sampling is followed, for example, by Elmo Roper (11) in his surveys for *Fortune* and by George Gallup (6) in his American Institute of Public Opinion surveys. The device, as has already been shown,

makes possible a more representative sample by imposing restrictions on the investigator's selection of cases for interview.

The technique of stratification, however, has its disadvantages. They arise from two sources: one, that relationships have been established (by prior test) between the variable under study and the different strata; and two, that the ratio or proportion each stratum bears to the total universe has been determined. The first condition is not difficult to handle. Simple experimentation will expose the most obvious relationships and the more subtle can be determined by careful cross-tabulation. The second—correct proportions within each stratum—is troublesome. It means virtually that one must know the proper allocations for any plan of stratification to be followed.

When the A. C. Nielson Company set out to put its British Food and Drug Index into operation in England, it first had to conduct a *census* of food and drug stores in order to know how many small, medium and large stores to allocate to its sample of stores. This, of course, was because England has no census of business such as we know in the United States. The Nielsen experience demonstrates the first step in stratification, although in many cases estimates of the total instead of a census would suffice.

The *U. S. Census of Population* (14) is useful for many groupings, such as age, sex, size of market, geographic area, race, etc. But it does not reveal the allocations of families or individuals by income groups. In this field there is considerable disagreement. For example, two independent research organizations set out to measure the audiences of one evening of broadcasting in Hartford, Connecticut (5). Below are the "controls" established by each organization.

DIFFERENCES IN SAMPLE DISTRIBUTION BY ECONOMIC LEVELS

Crossley (Pure Recall)		Hooper-Holmes (Aided Recall)	
Median Rentals	Distribution	Rental Values	Distribution
$54.00 . . . . Class A	26.7%	Over $83.33	Class A 20.9%
$40.00 . . . . Class B	36.7	$50.00-$83.32	Class B 19.8
$33.00 . . . . Class C	30.0	$41.60-$49.99	Class C 17.9
$25.00 . . . . Class D	6.6	Under $41.60	Class D 41.4

Even though the actual field techniques for obtaining the audience data had been the same, there is reason to assume the differences in sampling would have contributed significant variations to the final results.

Roper, in his nationwide sampling, maintains six controls: age, sex, population groups, geographic areas, occupation and economic levels. In Roper's sampling by economic levels, he takes into account "geographical variations in average income, as well as the size of place variations in average income." The Roper allocation of interviews by income groups, for the country as a whole, is as follows: A ("prosperous") = 7 per cent; B ("upper middle class") = 25 per cent; C ("lower middle class") = 45 per cent; and D ("poor people") = 23 per cent (10).

The "controlled" or "stratified" method, even though exact proportions for each stratum are not always available, is today's best known method for insuring a representative sample.

CHECKING THE REPRESENTATIVENESS OF THE SAMPLE

After a survey employing any one of the three sampling techniques (*accidental, random* or *stratified*) has been tabulated, there are ways in which the sample's representativeness can be ascertained. In each interview or mail return, provision is usually (should always be) made for the inclusion of certain identification factors, such as age and sex of respondent; does his family own or rent; does the family have any of the following: automobile, telephone, electricity, radio, refrigerator, gas, and the like.

On most of these factors fairly exact data are available by counties and in summary by many different breakdowns. By tabulating for these factors the sample can be checked against the known proportions. The assumption is that if the sample checks on items for which accurate information is available, the findings on the other questions will be representative.

One further consideration regarding representativeness, which is equally important in all types of sampling, concerns the "not-at-home" cases. If the problem under study is related to age or size of family, for example, missing "not-at-homes" introduces a serious bias. Studies have shown that the "not-at-home" family is usually younger and smaller than the average. Therefore, if the survey involves leisure time activities (such as reading or radio listening) or consumption of foodstuffs, the sample would not yield representative results.

In order to safeguard the sample's representative quality, each case must be accounted for. This applies specifically to random sampling as used in telephone surveys[2] and mail studies. In the latter type of survey the fact that the original mailing may be a random sample of the universe is of little consequence if, through self-selection, the responses do not represent the outgoing list.

IS THE SAMPLE LARGE ENOUGH?

Experience in the past decade has demonstrated — both in the biological laboratory and in the field of consumer research—that the number of observations or interviews (size of sample) is not nearly so important as the sample's representativeness. Therefore, the trend in the past several years has been toward smaller and more typical samples.

While there are still those who are impressed by large samples, the advanced research worker is not concerned with size *per se*, but rather with the interpretation of his findings in the light of the sample's size. Here the contributions of Brown have been exceedingly helpful in the field of market research (3, 2).[3] Brown has computed a table showing the size of sample required to yield various degrees of statistical accuracy. The size of the total sample necessary for a specific study depends on the sub-groups, or breakdowns, to be made and the expected proportions of the answers, as well as the degree of precision desired. For example, a much larger sample is needed to insure specific limits of statistical accuracy when the answers come out 50 per cent "yes" and 50 per cent "no" than when the response splits 10 per cent to 90 per cent.

In planning a sample, if one can estimate approximately how the critical percentages will fall, i.e., within what

[2] In the coincidental telephone method, it is assumed that "non-answers" are not at home and, therefore, not listening at home.

[3] Brown's table is also available in *The Technique of Marketing Research* (2, pp. 396-402). New York: McGraw-Hill Book Co., Inc., 1937.

range, then the sample size can be determined to within the degree of precision required. When in doubt as to the outcome the recommended procedure is to plan the quota for a 50 per cent-50 per cent split on the important question.

While it is true that the larger the sample the smaller is the statistical error, it does not follow that a large number of cases reduces the sample bias. In the matter of numbers of cases the *Literary Digest* was safe; the difficulty resulted from a biased or non-representative sample. In considering the margin of error in any market research study three sources of error must be considered separately: 1) sample bias, 2) sample fluctuations, and 3) method of obtaining the information. This discussion has been confined to the first and second points only, although all three must be examined in accepting the findings of a specific survey.

The size of a sample becomes particularly significant in the interpretation of results. The application of the conventional standard error formula[4] will establish the approximate limits of statistical accuracy of the findings.[5] In many cases, because the sample is small, the differences between two percentage figures are not statistically significant. In other words, the differences may be due to chance, and hence are not "real."

This is one point where little has been done to educate the non-statistically trained business man. While he is quick to appreciate the need for representative sampling—the *Literary Digest* did much to bring this about—he does not grasp readily the theory of random or chance errors and, hence, accepts the difference between any two figures as representing true differences.

Examples of such lack of insight are abundant in connection with the interpretation given to small changes in radio program ratings published by the Cooperative Analysis of Broadcasting (CAB). Here, probably more than any place in market research, differences are considered *per se* without regard to sample size. For example, an evening program broadcast once each week, and checked in each CAB checking city (7), is listed as having an audience of 20 per cent of the radio families. That "rating" of 20 per cent can fluctuate almost 3.0 points in the next report and still be within the range of a chance error due to sample size. Or to put it differently, it would be unsafe statistically to make definite conclusions regarding a change in a program with a "CAB rating of 20" unless the change exceeded 3.0 points above or below 20.0 per cent.

The lack of consideration for sampling errors by the business man—and, for that matter, the lack of application of such measurements by the research man himself—is one of the problems in present day market research that is crying for solution. As time goes on, the problem may correct itself through trial and error, but until it is understood there will be much unnecessary concern

[4] Discussed in Brown's paper (*supra*). It should be pointed out, however, that this is an old formula and is being rapidly discarded for newer small-sample statistics which emphasize the variance in the estimates of the population parameters rather than the particular estimate under consideration. By using these newer methods it is possible to determine the significance of a split at the 50-50 level with the same degree of accuracy that the traditional methods yield only when the split is more marked. For discussions of these newer concepts see Fisher (4) or Lindquist (8).

[5] Some researchers check the adequacy of a sample's total size by successively cumulating a number of small sub-samples until a stabilized value is reached. The assumption is that the sample's overall size is adequate when such a point is attained, i. .e., where the addition of more cases does not alter the results significantly. Today's trend is away from this rule of thumb procedure. For a discussion of the plan *see* Brown (2).

over differences that are insignificant and would, in all probability, vanish if the sample were enlarged.

SUMMARY

In summary, the purpose of sampling in market research, as in other fields such as chemistry, biology, and psychology, is to obtain data about the universe under observation with a certain degree of precision and with minimum expenditures in time and money. It is less expensive and quicker to sample than to take a census. Taking a sample involves two important considerations — *representativeness* and *size*. Each is subject to check and control. Adequate representation in the sample can be best assured through the method of "stratified" or "controlled" sampling. Proper size can be determined in relation to the required precision of the answers by the application of the conventional standard error formula reduced to tabular form by Brown.

REFERENCES

1. American Marketing Association. *The Technique of Marketing Research.* New York: McGraw-Hill Book Co., Inc., 1937.
2. Brown, Theodore H. Size of a Sample in Market Surveys. *The National Marketing Review,* 1936, 1: 258-263.
3. ———. *The Use of Statistical Techniques in Certain Problems of Market Research.* Boston, Massachusetts: Harvard University Graduate School of Business Administration, Business Research Study 12, 1935.
4. Fisher, R. A. *Statistical Methods for Research Workers.* Edinburgh: Oliver & Boyd, 1936.
5. Hettinger, Herman S., King, Robert N. and Peter, Paul A. Roster, Coincidental, Unaided Recall: How They Compare in Counting Listeners. *Advertising and Selling,* 1940, 33 (August): 20-21.
6. Katz, Daniel and Cantril, Hadley. Public Opinion Polls. *Sociometry,* 1937, 1: 155-179.
7. Lehman, A. W. Program Popularity in 1940. *Broadcasting 1941 Year Book,* 20 (3-A): 24, 38.
8. Lindquist, E. F. *Statistical Analysis in Educational Research.* Boston: Houghton Mifflin, 1940.
9. McNemar, Quinn. Sampling in Psychological Research. *Psychological Bulletin,* 1940, 37: 331-365.
10. Roper, Elmo. Classifying Respondents by Economic Status. *The Public Opinion Quarterly,* 1940, 4: 270-272.
11. ———. Sampling Public Opinion. *Journal of the American Statistical Association,* 1940, 35: 325-334.
12. Stanton, Frank. Notes on the Validity of Mail Questionnaire Returns. *Journal of Applied Psychology,* 1939, 23: 95-104.
13. Stephan, Frederick F. Representative Sampling in Large-Scale Surveys. *Journal of the American Statistical Association,* 1939, 34: 343-352.
14. U.S. Census of Population. *Final Volumes, Census, 1930.* Volumes 1-6.
15. Wilks, S. S. Representative Sampling and Poll Reliability. *The Public Opinion Quarterly,* 1940, 4: 261-269.

CHARACTERISTICS OF THE QUESTION AS DETERMINANTS OF DEPENDABILITY

By JOHN G. JENKINS, PH.D.
UNIVERSITY OF MARYLAND

IT is a commonplace to say that the responses obtained are a function of the question that is asked. There are probably few psychologists who would hesitate to subscribe heartily to this statement, yet there are many who by actual practice suggest that they are rendering it only lip-service. From Binet and Stern down to the most recent publications in legal psychology, there is an abundance of evidence to show dramatically how the form, or the wording, or the placing of a question may change the distribution of responses from the same group of respondents. A classical article by Muscio (8) and a confirmation and extension by Burtt and Gaskill (2) will come at once to the mind of a psychologist confronted with the problem. The evidence they contain points to the marked effects of exchanging a definite for an indefinite article and of similar small changes in the wording of a question. Because the matter by no means ends there, the writer has been invited to review specifically some of the other influences which have been found to reduce the dependability of questionnaire responses.

Although classification may be attempted along a variety of lines, a review of available considerations of the subject suggests that there are four major ways in which questions may influence answers so as to reduce more or less sharply the dependability of the responses given. These are:

1. *By biasing or predetermining the answer.*
2. *By failing to determine sufficiently the direction of response.*
3. *By exceeding the ability of the respondent to answer accurately.*
4. *By inviting inaccurate responses.*

Each of these will be discussed and illustrated. Before this is done, a note of caution must be sounded. The expert in this field need not be reminded that many of the sources of dilution listed above can be detected only by pretesting: some of his less-experienced colleagues have shown by questionnaires they have fostered that this is necessary. Many of the illustrations used in this article will deal with shortcomings in questions which no a priori analysis could have revealed. These shortcomings become apparent—and, indeed, could become apparent—only when the questions are tested in the field on samplings of the population later to be studied. It must be realized that a verbal acquaintance with causes of decreased dependability is of little value unless it is applied by searching the results of actual pretests for indications that important rules have been contravened.

1. BIASING OR PREDETERMINING THE ANSWER

A. *Through leading questions.*—Legal practice and common social usage have so long decried the use of leading questions as to lay anyone who uses them

open to the suspicion that he is seeking a predetermined answer rather than a valid response. It seems necessary to warn against such questions only because they continue to appear in questionnaires which are apparently designed in good faith to lead to publishable data. The topic may be dismissed with the statement that there is never any justification for the use of the leading question in serious research.

B. *Through loaded questions.* — We have spoken of questions as "leading" when they have been deliberately phrased by the writer in an effort to secure a particular answer. This is really a special form of the *loaded question,* that is, the question in which some biasing influence increases the probability of response in a given direction. In commercial market research agencies, it is widely recognized that the letterhead on which a questionnaire is printed may "load" the questions toward or against the sponsor. For this reason it has become common practice to send out questionnaires on plain paper and with the return address of a "neutral" firm. A report, made available to the writer by a commercial market research group, shows statistically reliable differences in response-frequencies when the identical questionnaire was mailed to equated groups, half of the mailing on company letterhead and half on the letterhead of a magazine.

A second type of "loading" comes from the inclusion of words or names which have emotional coloring for certain of the respondents. It has been shown that a sampling of the population materially changed the proportion of affirmative response to a question regarding the support of a specific governmental policy when a second form of the question labelled it as a policy of the president then in office (11). It has also been shown that doctrines labeled as "fascistic" or "communistic" are less acceptable to the populations studied than are the same doctrines when presented without the labels. Anyone who is familiar with the propagandistic device of *name-calling* will recognize at once how deeply the use of particular names and words may affect the distribution of responses to a question.

It is also recognized that the alteration of a question by the insertion of a qualifying word may profoundly affect the responses it evokes. Data from the office of the Psychological Corporation (10) may be cited in evidence. The question, "Are you willing to have an increase in prices with the hope that it will bring back prosperity?" brought an affirmative response from 69.6 per cent of the population interviewed. When the statement was changed to read "a *reasonable* increase in prices," 81.0 per cent responded affirmatively.

C. *Through questions which force the answer into unwarranted categories.*— A standard manual of instructions for the construction of questionnaires states unequivocally: "Whenever possible, yes-no questions should be used." The writer has suggested elsewhere (5) that this should be revised, in the light of available evidence, to read: "Yes-no questions should never be used except when all that is desired is an unequivocal choice between two mutually exclusive alternatives." Ghiselli (3) has shown that the mere change from a two-level to a four-level response may alter not only the proportion but as well the actual sign of the response to a question. The sole virtue of the yes-no question in many contexts is that it is easy to tabulate. To use it in situations where a graded response is possible is to invite inaccuracy of response by forcing the respondent to select one of two extreme

positions.

The check list, again, is suspect in many of the situations in which it is employed. The writer (4) has presented evidence to show that, unless all of the possible responses have been determined by previous research and are represented in the list of choices, the result may be skewed beyond recognition. This view has gained confirmation in data communicated privately by the Market Research Division of the General Electric Company. In checking use of a group of magazines, they found that the elimination of the name of a heavily favored magazine from a check list resulted in a distorted table of preferences. It should be added that the inclusion of a space labeled "others" and the request to supply items not included on the check list does not appear to improve the accuracy of the results.

D. *By improper order of questions.*—A recent textbook on market research urges its readers to arrange their questions in *logical* order. By contrast, anyone who has worked extensively in this field would be impelled to reply that the *logical* order would prove frequently to be a poor *psychological* order and that the matter could be settled only by careful forethought and adequate pretesting. If a question suggests the answer to a later question or if it is found to induce in the respondent an attitude which may seriously affect the following question, the demands of accuracy require a revision of the order, no matter how "logical" its original position may appear to be. Difficulties of the question-order, it should be added, may creep undetected into any questionnaire: their detection and removal are among the important functions of pretesting.

2. FAILURE TO DETERMINE SUFFICIENTLY THE DIRECTION OF RESPONSE

A. *Through the use of vague or indefinite questions.*—Succinctly phrased, the point involved might read, "Vague questions can be expected to return answers of vague significance." To ask respondents "Do you approve of the President's foreign policy?" is to lump together our relations to all the countries of the world, belligerent and non-belligerent, large and small, neighboring and distant. One suspects that anyone who gave an unequivocal answer would be responding largely in terms of subcortical impulses. Yet such questions have been employed within the year.

B. *Through the use of ambiguous terms.*—It is obvious that, if a term is ambiguous, the responses will fall into two or more classes in a way that defies all later attempts to separate the classes without repeating the research. Experienced workers will agree that, in spite of one's best efforts, ambiguous terms do creep into questions where they are detected only through pretesting. "Do you own a *light* coat?", and "Do you use *matching* nail polish?" are two examples which have recently come to the writer's attention.

C. *By failing to pretest for tacit assumptions.*—Pointing out that the respondent never replies to the question-as-asked but always to the question-in-a-context, Lazarsfeld (7) has laid great stress on the necessity of detecting the tacit assumptions made by different respondents with regard to a given question. In normal social intercourse, one replies differently to the question, "Who will be at your house tonight?" depending upon whether the questioner is the grocer, the friend of ten-year old son, the veterinarian, or a prospective guest. Answers to the question, "What do you want in a new car?" may vary considerably as the individual respondent decides that this means the car actually planned

on, any new car in the appropriate price range, or an ideal car regardless of price. In a local pretest of a questionnaire, it was found, completely to our surprise, that certain negroes among the sampling took an ethnological view of the question, "Do you take many colored pictures?" Just as ambiguous words give unmeasured distortion to the distribution of replies, the existence of unrecognized tacit assumptions will reduce its dependability. In both cases adequate pretesting will serve in an important preventive capacity.

D. *By failing to provide for homogeneous categories of response.*—Lazarsfeld (6) has pointed out that, unless guidance is given, those who respond to certain types of questions may reply in terms of attributes, or in terms of influences, or in terms of tendencies. This, in itself, is harmful only because we have no way of being assured that those who named influences might not have been able to give significant replies in terms of attributes as well. Thus if 10 per cent say that they went to X University because of its traditions (attributes) and 35 per cent state that they came on the advice of friends (influence) we are treading on doubtful ground if we draw our conclusions directly from these results. There are two ways out. One is to separate the answers which speak in terms of influences from those which refer to attributes and to make comparisons only within the groups. The other involves employing an "elastic" questionnaire. Thus if a person speaks of advice of friends (influence), he is asked what advice they gave (probably attributes), while the person who mentions traditions is asked how he learned about the traditions. Either of these two ways involves more effort on the part of the director of the study but it reduces the probability of comparing non-homogeneous data in ways likely to produce a conclusion which inadequately reflects the actual situation.

3. EXCEEDING THE ABILITY OF THE RESPONDENT

A. *By the use of words not understood by the whole sampling.*—No psychologist needs to be reminded of the impoverished vocabulary which typifies the bulk of the population. Only by pretest in the case of specific questions, however, is he likely to grasp the dramatic limitations of popular vocabularies. Particularly dangerous are those words which sound familiar but are incorrectly interpreted. An unpublished study by the writer revealed that although almost all of the students investigated claimed acquaintance with the word "peer," the vast majority of them defined it as meaning "superior to." Poffenberger (9) found that a group of housewives had no preference between a "Ham for Epicures" and a "Ham for Igorots." Pretests not infrequently show that a keyword is not comprehended by a significant proportion of respondents and that some other term must be substituted.

B. *Through the use of complex questions.*—It would seem that a schoolboy would grasp the undesirability of using double-barrelled questions, yet one encounters them not infrequently in current questionnaires. It would be difficult to interpret a "yes" or a "no" answer to the question, "Are the professional opportunities and scale of salaries at your university satisfactory to you?" There may be situations which justify the deliberate use of two-questions-in-one, but the writer has not yet encountered one.

C. *By failing to aid the respondent to remember.* — Lazarsfeld (1) makes a very strong case for the desirability

of "interviewing along the time-line" of the original action as a way of aiding the respondent to remember the specific materials desired in a given questionnaire. This is, of course, an extension of the common-sense method of seeking to retrieve a lost article by memorially retracing one's steps from a time when it was still in one's possession. This requires the use of accessory questions, the answers to which may be unimportant in themselves, in order to lead the individual through the necessary temporal stages to the critical questions.

D. *By failing to aid the respondent to express himself.* — Particularly as one goes down the educational scale, respondents are inhibited by difficulties in expression. Weaver (12) has found it helpful to provide pictures of the choices with which he is concerned, allowing for checking by the artistically inept, but providing also a base of operations for those who can handle a pencil productively. Where words must be used with respondents of badly limited vocabularies, preliminary research to set up exhaustive check lists is justified by the elimination of any requirement that the respondent shall make ready use of unfamiliar verbal tools.

E. *By employing poorly chosen units of response.* — Questions are frequently seen which call for responses in units either difficult or impossible for the respondent to handle. There are many who could not answer the question, "How many pounds of coffee do you buy per year?" When this is phrased as "How many pounds of coffee did you buy last month?" the dependability of the response increases materially. If it is desired to extrapolate from this information to an estimate of annual consumption, provision may have to be made for some indication of what the respondent considers to be his normal monthly consumption. In any event, good practise demands that the units employed in any response should not exceed the expected ability of the respondent to employ them reliably.

4. INVITING INACCURATE RESPONSES

A. *Where an accurate response would embarrass the respondent.*—Only a pretest can show which questions are likely to embarrass the respondent sufficiently to tempt him into inaccurate responses. Social pressure to give a "good" response, together with local and personal inhibitions of various sorts, may constitute serious barriers to dependable answers. If embarrassment is disclosed by a pre-test, the topic must be approached indirectly or studied in some fashion that avoids the face-to-face inquiry. In one case, for example, it was found that front-door inquiries regarding magazines read in the home brought an undue preponderance of the "better" magazines, while tabulation of magazines collected for charity from the back-door restored the expected percentages of wood-pulp literature. In some cases, indirection has taken the form of inquiring not about the respondent but about his neighbors. In any event, one is certainly not justified in expecting universally honest attempts to return accurate answers to embarrasing questions.

B. *Where an accurate response would incriminate the respondent.* — Certain types of inquiries lead into fields where accurate responses to particular questions might actually incriminate those who reply in specified ways. There seems to be nothing to do except to abandon such questions when they are detected either in the original make up or in the pretest. To ask a penitentiary inmate to tell of all his earlier offenses contains a high element of futility—yet it has been done. To ask relief workers

to confess their records of illicit employment is to invite a meretricious response—yet it has been done. Questions of this type can serve only to bring about further distrust of the questionnaire method as a whole.

C. *Where an accurate response is for any reason inexpedient.* — Occasionally one finds questions which appear innocent but which, for local reasons, are inexpedient for the actual respondents to answer accurately. Where pretest or field use reveals this state of affairs, the question—at least in its direct form—must be abandoned. A question on attitude toward unions which had been effective elsewhere ran into substantial blocking when it was attempted in a city which had just witnessed the end of a sanguinary strike. Another question, dealing with the use of metal containers, was regarded with high suspicion when it was asked in a city containing a high percentage of disgruntled employees of glass factories. It would be optimistic beyond justification to expect dependable replies in such cases, and one must either make an indirect approach or give up the questions involved in obtaining an accurate response.

CONCLUSION

The gap between the best possible questionnaire and many of the specimens currently circulated is enormous. At best, this instrument has serious shortcomings against which may be balanced the fact that certain topics can be investigated only by questionnaire. Where as many as possible of the diluting influences have been eliminated, where a properly prepared questionnaire is distributed to an adequate sampling, and where results are interpreted in the light of careful analysis, the questionnaire ranks high both in utility and in dependability. Those who wish to use it at this level will do well to keep constantly in mind a list of major sources of undependability in the form, placing, and content of their questions.

REFERENCES

1. American Marketing Society. *The Technique of Marketing Research.* New York, McGraw-Hill Book Co., 1937, 64.
2. Burtt, H. E. and Gaskill, H. V. Effect of the Form of a Question. *Journal of Applied Psychology,* 1932, 16, 358.
3. Ghiselli, E. E. All or None versus Graded Response Questionnaires. *Journal of Applied Psychology,* 1939, 23, 405.
4. Jenkins, J. G. *Psychology in Business and Industry.* New York: J. Wiley & Son, 1935. Pp. 348.
5. ———. Questionnaire as a Research Instrument. *Transactions of the New York Academy of Sciences,* 1940, II, 2, No. 5.
6. Lazarsfeld, P. F. Psychological Aspect of Market Research. *Harvard Business Review,* 1934, 13, 54.
7. ———. The Art of Asking Why in Marketing Research. *National Marketing Review,* 1935, 1, 1.
8. Muscio, B. Influence of the Form of a Question. *British Journal of Psychology,* 1916, 8, 351.
9. Poffenberger, A. T. *Psychology in Advertising.* 2nd Edition. New York: McGraw-Hill Book Co., 1932. Pp. 346.
10. Roslow, S. and Blankenship, A. B. Phrasing the Question in Consumer Research. *Journal of Applied Psychology,* 1939, 23, 612.
11. Roslow, S., Wulfeck, W. H., and Corby, P. G. Consumer and Opinion Research: Experimental Studies on the Form of the Question. *Journal of Applied Psychology,* 1940, 24, 334.
12. Weaver, H. G., *Proving Ground of Public Opinion.* Detroit: General Motors Co., 1940.

EVALUATING THE EFFECTIVENESS OF ADVERTISING BY DIRECT INTERVIEWS

By PAUL F. LAZARSFELD, PH.D.
COLUMBIA UNIVERSITY

THERE has been much discussion among phychologists about whether there is any point in asking people why they buy, or why they do certain other things with which a market research study might be concerned. The problem in all its generality cannot be treated in a short article. One aspect of it has had special prominence in market studies which attempt to test the effect of advertising. Is there any chance, by means of direct interviews of evaluating the role played by advertising in a specific group of purchases? This paper is restricted to this aspect of the problem and to a clarification of the following questions involved:

1. What does it mean to speak of advertising as having played a role in a certain purchase?
2. What kind of interviewing technique is likely to bring out pertinent information?
3. What meaning does a statistical analysis of such reasons have?
4. How can results in this field be verified?

1. *What does a "reason for buying" mean?*—The meaning of the word, "reason," just as with most other concepts having a causal connotation, must be derived from the idea of a controlled experiment. Let us assume that we are dealing with a group of people who have been divided into two groups by a completely random procedure, so that no single characteristic will give any individual more of a chance of belonging to one, rather than to the other group. Next, one of these two groups is exposed to certain advertising, but the other is not. Then, after some time has passed, if the exposed group buys more of the advertised product than the unexposed group, we can say that the advertisement has caused increased sales. This, of course, is a definition, and other students may prefer to use other starting points for this kind of discussion. As long as we do not deviate from this definition through the discussion, however, it will not be difficult to relate other approaches. It should be kept in mind that we are dealing with two groups which have been made comparable by random selection and subsequent exposure; this is quite different from comparing two groups who have or have not been exposed to advertising by their own choice.

The simplest scheme for presenting the relationship between exposure and effect is the fourfold table. Table I gives a schematic example of this kind of de-

TABLE I

RELATIONSHIP BETWEEN EXPOSURE AND EFFECT IN ADVERTISING

	Bought	Did Not Buy Product	Total
Mail was sent . . .	500	500	1000
Mail was not sent	250	750	1000
Total	750	1250	2000

signed experiment, assuming that 2,000 people have been divided at random into two groups of 1,000 each; and that one of these groups has been sent a mailed advertisement; while the other has not. If after some time it is found that 500 members of the exposed group now own the product advertised by this mailed circular and only 250 members of the unexposed group have bought it, the mailed advertisement can be considered the cause of the difference in the purchases of the product between the two groups.[1]

Such an effect can be measured with all the various coefficients which have been developed to gauge the correlation of a fourfold table.[2] All these measurements go back to one simple idea. If the exposure being studied had not had any effect, then the exposed and the unexposed people would have made the same number of purchases. Of the group unexposed to the advertising, 25 per cent bought the product. Twenty-five per cent of the exposed group could then also have been expected to buy the product without the instigation of the advertisement. But actually 50 per cent of the exposed group did so. The 25 per cent difference, therefore, can be attributed to the advertising. To speak of persons, not per cents, it can then be said that 250 of the exposed group who bought the product would have bought it anyway. The other 250 bought it "because" of the advertising campaign. From now on these 250 cases will be referred to as "*b*-cases." The 25 per cent buying difference between the exposed and the unexposed group will be used as a measure of the effect, *F*, of this special campaign.

Still referring to this imaginary example, we originate the psychological problem if we ask the following question: *By means of direct interviews, is it possible to single out these 250 b-cases?* So far we know only that among those people who were exposed and bought the product, there are these 250 *b*-cases. But, from the statistical analysis, we cannot say whether it was Mrs. Smith or Mr. Jones who belongs in this group. Would they be able to tell us? As far as we can see, this is the only way to formulate the problem of whether people know why they buy: If a controlled experiment, of the kind described above, were made, would we be able to single out, by means of appropriate interviews, these 250 *b*-cases which we find by statistical cross-analysis?

The question can be put in still another form, by giving the number of *b*-cases a slightly different interpretation. The correlation between exposure to the mailed advertisement and buying is by no means perfect. As a matter of fact, in practical work, even if it is at all significantly above zero, we should consider the campaign quite successful. The correlation is not high because it does not make much difference to the buying habits of most of the people whether or not they receive the advertisement. On the other hand, for a few people, it makes all the difference in the world. They would not have bought this merchandise if the advertisement had not reached them. Theoretically speaking, then, the people we are studying were of two kinds. Some find themselves in such situations, or in such moods, or have such personal characteristics, that there is a correlation of +1 between their buying and their exposure to our advertising campaign. For the

[1] This discussion, which is designed to lead to psychological problems, omits the question of statistical reliability.

[2] G. U. Yule and M. G. Kendall. *An Introduction to the Theory of Statistics.* London: J. B. Lippincott Co., twelfth revised edition, 1937. Pp. 34 ff.

rest of them the correlation is zero, indicating that they would or would not have bought, regardless of their exposure. It could be shown algebraically, but it is obvious even by meditating about the matter, that the number of *b*-cases is simply the number of people in our exposed groups who have those characteristics pertinent to a correlation of +1 between exposure and action.

However one looks at it, then, the question of whether a certain advertisement was a reason for a person's buying a commodity is equivalent to a question such as this: When the person was exposed to this influence, was the situation such that without this exposure he would not have bought the article, whereas since he was exposed his buying necessarily followed according to all the "rules of psychology" we were able to apply to this case?

2. *Interviewing techniques to detect the role of advertising.*—There are certain interviewing techniques which can be derived as necessary from the preceding analysis. According to it, the art of asking why consists in getting enough information about the situation in which the respondent was exposed to advertising so that a reasonable estimate can be made of whether or not he would have bought the advertised product even if he had not been exposed. We put ourselves, so to speak, back to the time of the exposure, and try to forecast in each single case what would have happened. This point of view allows a scrutiny of the actual information to determine whether it is complete enough to permit such a prediction.

A friend comes to visit me late at night, and when asked why he has come, he answers, because he "feels lonely." Putting ourselves in his situation at home when he felt lonely, we see immediately that from that point we could not have predicted his visit to me if we did not have any further information of this sort: Does he have other friends as well? If so, why is the visit to me more suited to a feeling of loneliness than a visit to someone else? What is it in his personality or in the quality of his feeling of loneliness which makes him want to see a friend rather than go to a movie or get drunk? It should be observed that we do not want to know how these things developed biographically. We need merely additional information at the time of his loneliness to "understand," or to have predicted for this phase of the development, that in the end he would come to visit me.

More examples of this kind can now be treated with fewer of the details.

A respondent says he didn't vote because it rained on election day. That might or might not be an acceptable statement, according to the following additional information which would be needed. Is his health frail so that he is always afraid of going out in the rain, and therefore the rain actually would have led us to predict his not voting? Or is he just not interested in politics and any other excuse would have been as good? Or was the rain so bad that, according to prevailing standards, no one without a special means of transportation would be expected to go out on the street?

A housewife uses a brand of soap she has not used before. In an interview she reports that she could not get her usual brand, and so bought the one which she found at her dealer's. Whether availability at the dealer is to be noted as a reason depends mainly upon the question of whether the new soap was really the only one which the dealer carried at the time of her purchase. If this is not the case, then, had we observed her at the store, we would need some more information for prediction: It might have been the only soap the name of which was known to her from advertisements; or the only soap available within a certain price-range; or the soap displayed nearest to where she stood in the store. Questions directed toward such information would be indispensable in making the interview complete.

A man says he votes for the Democratic Party because Roosevelt is for the poor fellow. But so are the Townsendites and the Communists. We therefore need additional information of this kind: Does the man know of other parties? Does he think that a vote for small parties is a wasted vote? Has he examined the other parties and found shortcomings in them which negate their attitude toward the poor fellow? Again, only with such additional information can we decide the weight which Roosevelt's social policies have had for this man's vote. It might well be that in the course of such a scrutiny, the whole picture would change, and other reasons would emerge into the foreground.

Even an answer which is usually quite meaningless and comes only from bad interviewers can make sense with this sort of examination. Suppose we study why people choose certains items on a menu, and someone tells us that he chose shrimp salad because he likes it. This answer might be meaningful if it is qualified, upon further questioning, somewhat like this: He liked shrimp better than anything else on the menu that day; or he liked it best in the price-range he was prepared for; or there are other things on the menu he likes as well, but he had eaten these foods recently.

In order to work such techniques into a formal questionnaire to be used for a specific problem, it is usually necessary first to make trial interviews and to scrutinize the way the consumer, himself, looks at the situation. Taking as an example a questionnaire on gasoline purchases, it was found for instance that when people were asked how they started to use their present brand, they usually gave one of five reasons. Either they spoke (1) about their personal reasons, such as relationship to station attendants, (2) about convenient location, (3) they just tried around and then got used to a brand, (4) they mentioned recommendation of friends, or (5) finally advertising. The research problem then consists in devising such a questionnaire that no matter what starting point the respondent takes of his own initiative, provisions are made to check back to see whether additional factors played a role which the respondent overlooked under the impact of the first factor which came to his mind. The pertinent part of such a questionnaire, then, would consist first of a general question, "Why?" and a check list of the reasons which can be expected on the basis of the pretest. Then the list of *specifying questions* would be added. From the schedule used in the gasoline study we present the specifying questions as an example (without regard to any particular spacing).

IF—(1): *Personal Reasons (Friend or relative in station, etc.)*
- *a*) What sort of connection do you have with the attendant?
- *b*) Do you care what gasoline you buy provided it is a known brand?
- *c*) (If yes to *b*) What is your preferred brand?
- *d*) (If *c* is answered) How did you happen to get interested in this brand?

IF—(2): *Convenience*
- *a*) Are other brands as conveniently located?
- *b*) (If yes to *a*) Why then did you select this brand?
- *c*) (If no to *a*) Why is there just this one?
- *d*) Are there many stations with your brand thus, making it convenient?
- *e*) What other advantages besides convenience has this station (or brand)?

IF—(3): *Trial*
- *a*) (If chance) How did you happen to try it?
- *b*) (If deliberate) What other brands did you try?
- *c*) (If deliberate) About how long did you experiment?
- *d*) (If deliberate) What made you finally decide in favor of this brand?

IF—(4): *Recommendation of friends or acquaintances*
- *a*) Who was the adviser (i.e., personal friend, relative, business ac-

quaintance)?
b) Did he tell you about the product or did you just follow his example?
c) (If the former—*b*) What argument made you follow his advice?

IF—(5): *Advertising*
a) Was it just one advertisement or a series of advertisements?
b) Just what advertisement was it that impressed you most?
c) Can you remember when you saw or heard it?
d) (If radio—do not ask) Where was it?
e) What were the sales points that impressed you most?
f) Can you give more details as to how the advertisement impressed you and made you buy this brand?

After these interviews are turned in, it is then usually possible, with a great degree of safety, to divide those cases where it seems that advertising was a cause of a purchase from those where it does not seem to have been a cause. In one study, for example, five judges were used to make this decision on the same material. Among them an average tetrachoric correlation of .85 was obtained.[3]

3. *Statistical analysis of reasons.*—Suppose, then, that we find that 20 per cent of a number of people who bought the product acknowledged the effect of advertising. What does this figure mean?

In order to give the answer, we must go back to the original fourfold table, reproduced here in a somewhat more general form.

[3] Other examples of this interviewing technique can be found in two earlier papers by the writer, "The Art of Asking Why," *National Marketing Review*, Summer 1935, 1 (1):1-13, and "The Use of Detailed Interviews in Market Research," *Journal of Marketing*, 1937, July:3-8. In these papers, however, the technique of specifying questions was recommended only as useful, and was not linked up with an analysis of the concept of "reason," itself.

TABLE II

EFFECT OF ADVERTISING ON BUYING

	Bought Product	Did Not Buy Product	Total
Exposed	q	t	e
Not exposed to advertisement .	s	v	$n-e$
Total	a	$n-a$	n

One way to express the effectiveness, F, of exposure to advertising upon buying, would be the formula:

$$F=\frac{q}{e}-\frac{s}{n-e}$$

(Differences in proportion of buyers between exposed and unexposed people.)

The *number* of *b*-cases would be expressed by the formula:

$$b=e.F$$

where e equals the number of exposed people.

The number of *b*-cases, then,—that is to say, *the number of people who say they bought because they were exposed, equals the effectiveness, F, of the advertising, times the extent of exposure.* This is an important result. It shows that if we have interviews which are able to reconstruct the past so well that the information we have is equivalent to a controlled experiment, then the statistical result we get is the product of two factors: the extent and the effectiveness of an advertising campaign. The same number of *b*-cases might be found if very many persons are exposed to the advertisement even though its effectiveness is rather small, or if relatively few people are exposed to a campaign which is extremely effective, in the sense of getting exposed people to come to the store. The number of *b*-cases, then, measures what one might call the *impact* of an advertisement,

which is the product of its extent and its effectiveness.

To round out this analysis, it is necessary to visualize the three concrete research situations in which we are likely to use direct interviews. It can be that we have access only to people who would buy the product—as would be the case, for example, with a mail-order house, or when the respondents are already in the store, or when we are dealing with subscribers to a magazine. Using the symbols of Table II, this would be a situation where we have only the "a-people" who bought, but where we do not know how many people did not buy, and therefore cannot know the size, n, of the complete sample. In this case, interviews would tell us only the relative impact of two advertising media. For instance, if we have ascertained that b_1 people bought a product under the influence of the radio, and b_2 people bought it under the influence of the newspaper, then the ratio of b_1 to b_2 would be our main result.

The question must be raised whether such a result, even if it measured up to the highest standards, is of practical importance. There are a number of reasons why the answer should be "yes." It is true that the most interesting piece of information would be the effectiveness, F, itself. We should like to know what increases in sales a campaign brings about among those people who are exposed to the advertisement. Very often, however, it is enough for practical purposes to know only the impact. If an advertiser puts a certain amount of money into radio and another amount into newspaper advertising, he would like to know which medium got him more additional sales, and he wouldn't care whether a difference is due to the two media reaching the same number of people, but the one is more effective with those it reaches; or whether it is because the two media are equally effective but the one has a larger audience, or both of these. Also, in certain cases, additional material might permit an estimate of the exposure, e, for the two media. We might, for instance, know the approximate size of the audience to a certain program and the approximate readership of a magazine. If, then, an appropriate sample is interviewed and we find out how many buyers of the product have been influenced by the program and how many by the magazine, we can take into account the amount of exposure for the total population from which the sample has been taken and then compute directly the effectiveness, F, of the two media. F would tell us what increase of sales the medium would lead to in a controlled experiment.

If the research situation is such that we know how many people in our complete sample bought the product, and also know how many have been exposed to the different advertisements, we can immediately construct Table II, and then it would seem that no direct interviews would be needed. This, however, is not always correct. Even where a complete fourfold table can be drawn up, the direct interviews may add decisively to our knowledge. Results have been reached, for example, showing that people who are exposed to advertisements also use more of the advertised products.[4] In these studies we do not know to what extent exposure leads to use, or use to exposure. One way to disentangle the two factors is, again, the direct interview. The same is true in regard to spurious factors. People might listen to a program *and* buy the product because

[4] See the report of a study done by Cornelius DuBois, Continuity Study, *Tide Magazine*, 1940, 14 (21, November 1) :13.

they are well-to-do. Direct interviews would lead to a number of *b*-cases smaller than could be expected from a fourfold table. A spurious correlation would thus be corrected and wrong interpretation avoided.

4. *How to test the validity of reasons.* —So far we have assumed that our interviews yield results equivalent to a controlled experiment, and our concern has been to clarify the relationship between the statistical results obtained by the two methods. Can we expect, however, that direct interviews will ever reflect the process of a concrete purchase so clearly that a decision can be made as to whether a certain piece of advertising was or was not an influence? Much has been said about people not knowing why they act, and many a market research man feels that it is an impossible task to find this out. Here again, the schematic analysis of the problem presented here will help to clarify the matter.

First, it must be remembered that we are not dealing with the rather meaningless general question of why people buy, but with the task of tracing the effect of specific advertising upon concrete purchases with the help of retrospective interviews. In such cases, where do the difficulties lie?

In the first place, it may be that people forget that they have been exposed to advertising. The interviewing technique sketched above tries to follow possible exposure into situations where the respondent might have overlooked mentioning it. Whatever the respondents mention — role of price, or influence of salesperson, or special qualities of the product, we check back to see whether advertising entered by forming their idea of the quality, or giving advice of the salesperson additional support. This should reduce mistakes in remembering exposure considerably, probably leaving only those cases where such exposure would not have had any actual influence, and thus the omission would not change the results in any case.

In those cases where exposure has been established, errors can be made in appraising the effect. We shall certainly be frequently wrong in making the statement that, in a given situation, buying was or was not to be expected, because the phychological rules which we more or less explicitly apply are not very safe. But there still might be a chance of our erring sometimes on the positive and sometimes on the negative side, and that in the end the misjudgments would probably cancel out and the statistical results would be valid.

As a matter of fact, the only conceivable test of the validity of reasons is a statistical one. For obvious reasons, it is not possible to verify whether a certain purchase would have been made even without a specific advertising influence having been introduced. We can test only whether the number of *b*-cases which we ascertain by direct interviews is corroborated with the number of *b*-cases derived from a controlled experiment. This innocent little sentence implies a vast research program. In order to settle whether reasons for buying obtained in an interview are valid data, it would be necessary to set up controlled experiments and parallel them by direct interviews. Opportunities are not lacking. Quite frequently advertising campaigns are tested by comparing two trade areas, one of which is covered by the campaign, and the other of which is not.[5] In each of these situations, direct interviews could be added and their results compared with the objective test.

[5] *See* Frank Stanton, A Two-Way Check on the Sales Influence of a Specific Radio Program. *Journal of Applied Psychology*, 1940, 23 (6, December) :665-672.

Such opportunities have not been used so far and one of the tasks of applied psychologists is to stress the desirability of this procedure whenever they are connected with a pertinent test situation. At present, therefore, only crude substitutes for such tests can be reported and used for exemplification.

In a market study, the problem was to find out why people do not buy at a certain department store. One reason to be tested was the location of the store. People were divided into two groups, according to whether they lived in the rather immediate neighborhood of the store, or in a more distant area which was defined by several criteria such as number of miles and transportation facilities. As a result of cross-analysis, a fourfold table of the following kind was obtained.

TABLE III

REGULARITY OF BUYING IN RELATION TO DISTANCE OF STORE

	Buy Regularly	Do Not Buy Regularly	Total
Live near the store . .	277	116	393
Live at some distance from store . .	135	302	437
Total	412	418	830

From a computation identical to the one given in Section 1 above, we can conclude that there should be 158 *b*-cases,[6] that is to say people who, upon being interviewed, would reply that they do not buy at this store because it is located inconveniently for them. The 418 people who did not buy at the store or did not buy there regularly were interviewed, and 40 per cent, or 167, said that their reason for not buying was the inconvenient location of the store. This is indeed a corroboration of the expectation derived from the fourfold table, which can be considered relatively satisfactory in view of the crudeness of the data.[7]

Another reason investigated in this study was: poor opinion of the store. In regard to this poor opinion a somewhat different kind of test was made. Somewhere during the interview each respondent who only bought occasionally at the store was asked to judge it according to the following seven aspects: style of merchandise, interior arrangement, salesclerks, price of merchandise, quality of merchandise, services, customers.

The answers to these questions can be used as a crude attitude-test by classifying people according to the number of items on which they express criticisms. The most favorable attitude would be that of the respondent who had no criticism to make on any aspect. The least favorable attitude would be exhibited by respondents who had criticisms on all these items. In classifying people according to this test, we perform an operation comparable to the one we performed when we classified people according to location. The score permits people to be grouped according to their critical attitude toward the store, and this grouping can be used for the purpose of studying the validity of people's

[6] $\left(\frac{277}{393} - \frac{135}{437}\right).393 = 158$

[7] As the study was not originally devised for the purpose for which it is being used here, the following tabulation of reasons was decided upon. Those people were considered most strongly influenced by location of the store who, as reason for not buying at the store, only mentioned its inconvenient location (33 per cent). Fifteen per cent said inconvenience and poor opinion of the store kept them from buying there. As there were no further data available, half the people were counted as influenced by inconvenient location. In a study more pointed to the problem of reasons, it would have been necessary to continue the interview and find out whether they bought at other stores more inconveniently located, whether they had less good opinion about other stores where they still bought, and so on, using the technique elaborated in Section 2.

reasons. We would expect that the more critical groups would more frequently mention bad opinion of the store as their reason for not buying there. The figures bear out this expectation as shown in Table IV.

TABLE IV

RELATION BETWEEN EXPRESSED CRITICISMS AND REASONS FOR NOT BUYING

Criticism score (Number of items criticized)	Total number of cases	Proportion of respondents giving bad opinion of store as reason for not buying
None	55	25%
One	45	48
Two	29	64
Three or more .	47	78

We see that the number of criticisms increases in proportion to the number of people who give poor opinion of the store as their reason for shopping there only occasionally.[8]

It has to be kept clearly in mind that neither of these two examples contains any circuitous reasoning, as one might assume at first glance. The information on residence and attitude toward the store is collected independently of the quest for reasons. Logically, it might be that residence and attitude, on the one hand, and reasons on the other hand, have no relationship. That they actually do is an inferential test for the validity of the reasons. Furthermore, it is important to realize that such factors as where a person lives and the attitude he has, have the logical character of an "exposure;" they take on the character of "reasons" if in an individual case they are linked up with not buying at the store on the ground of the respondent's retrospection and the psychologist's evaluation of it.

It seems from this and a few other available examples,[9] that using personal interviews to test the effect of advertising it not impossible. On few points could a psychologist be of more service to the men of action than by refining this type of interview and studying the extent and the limits of its applicability. The entire complex of problems, however, of which we have been able to present only a small segment, is difficult and will require a wide diversity of interconnected studies, beyond the working capacity of a single psychologist. Its obvious practical importance might warrant the coordinated efforts of applied psychologists as a whole.

[8] It is revealing to study in great detail the "deviate cases" in this table. One-fourth of the respondents do not voice any criticism in the attitude test but still give bad opinion of the store as their reason for not buying. This is easily explained by the fact that our simple scale worked only with six criticisms, whereas they by no means exhaust all possible objections which might be raised. An inspection of the cases in the first line of Table IV shows indeed that those women refer in their reasons to other items than the one about which they have explicitly been asked in the attitude questions. At the other end, there are 22 per cent of the people who have more than three criticisms but who do not give bad opinion as the reason for their not buying. These people, upon closer scrutiny, appear to be those for whom inconvenient location is also so dominant a reason that they would not buy at the store even if they did have a good opinion of it. Therefore, even if they have a bad opinion and do not buy there, the two factors are not causally connected.

[9] *See* E. Smith and E. A. Suchman. Do People Know Why They Buy? *Journal of Applied Psychology*, 1940, 24 (6, December) :673-684.

EFFECTS OF REPEATED INTERVIEWING ON THE RESPONDENT'S ANSWERS

BY FLOYD L. RUCH, PH.D.
UNIVERSITY OF SOUTHERN CALIFORNIA

EVER-INCREASING use of the personal interview, the mailed questionnaire, and even the pencil-and-paper personality test has led to more and more frequent demands upon the leisure time of Americans and more insistent intrusion into their private lives. It is reported, for example, that some of the housewives of Atlanta, Georgia, keep a week-by-week tally of the number of times they are interviewed. Perhaps General Johnson could spare himself the fruitless effort that was Diogenes's by taking a trip to Atlanta. At any rate it becomes increasingly important to concern ourselves with studying the effects of repeated interviewing on the answers we obtain.

The title of this short paper is to be interpreted as including any instance of repeated or lengthy questioning whether by personal interview or by mail, with the same or different questions, and/or by the same or a different research organization.

The earliest study on the effects of repeated interviewing on the respondents' reactions is that of Hollingworth (3). This worker had a group of ten subjects rate the degree of funniness of thirty-nine jokes which had had little circulation prior to the experiment and were presumed to be fresh experiences to the subjects. The rating or "interview" procedure consisted in having the subjects cast the jokes (which were printed on cards) into piles to indicate their varying degrees of funniness. The procedure was repeated on five different sittings at weekly intervals. The results show conclusively that certain of the jokes increased in their rated funniness relative to the group as a whole and that others declined.

Hollingworth's experiment makes it quite clear that repeated interviewing on the funniness of humorous situations markedly alters the respondent's estimates of their funniness. This laboratory study has been borne out by the experience of a commercial research organization in studying the respondents' preference for a food. In the first series of observations the respondents were permitted to choose from a series of unlabeled packages the food which most suited their taste. The standing of each brand was measured in terms of the percentage of the persons who selected it. For several weeks the product being tested rose in popularity relative to the other brands. Finally the new brand declined in popularity until it found a stable position in the hierarchy of preferences. This effect is consistent with that reported by Hollingworth and shows that the results of repeated interviewing or observation will not agree with those obtained in a single study placed at any one point in time.

It must be noted that the interview situation in the two examples given was such as to involve actual experience with the material concerning which the

subjects were being interviewed. Since in neither instance was there a control group which experienced the material in the absence of interviews concerning it, we are unable to differentiate the effects of the experience from the effects of the interview.

In a study of the influence of familiarization on preference Maslow (6) projected lantern slides of paintings on a screen and asked his subjects to look at them carefully. Each painting was exposed for a period of fifteen seconds on four successive evenings. Six days following the fourth showing the subjects were shown the pictures they had seen previously mixed with paintings of equal merit and by the same artists but not previously shown. The subjects rated the paintings on the basis of liking for them and their beauty. The familiarized paintings were liked much better than the unfamiliar ones, receiving and average of 18.3 points whereas the unfamiliar paintings received an average score of 5.9 points.

These studies make it quite clear that tastes in three fields are greatly influenced by familiarity whether this familiarization be a part of the interview situation or introduced independently.

The effects on the respondent of overlong interviews may be regarded as somewhat similar to those of repeated interviewing. Various observations show that an interview ballot which is too long bores and irritates the respondent and encourages error both in the interviewer and in the respondent. It is particularly dangerous to compare the data on two similar questions when one is part of a short ballot and the other comes at the end of a long one. This point has been made very effectively by H. C. Link (5).

During 1931 a survey of reader interest sponsored by *Liberty Magazine* was made in which the attempt was to measure the interest value of the contents of the *Saturday Evening Post, Liberty, Collier's* and the *Literary Digest*. This nationwide study created considerable comment as its results indicated *Liberty* to be far superior to the other magazines. For example, the results indicated that the readers of *Liberty* as compared with those of the *Saturday Evening Post* read and remembered much more of the editorial content and advertising copy. Even the *same* advertisement appeared to be better read and remembered when in the smaller magazine.

Link indicated an important defect in the technique which might well have been responsible for the superiority of the same advertisement in *Liberty* as compared with the *Saturday Evening Post*. The interviewers ran through whole copies of the magazines with the respondent to find out from the latter which elements in each had been read and remembered. Since the *Saturday Evening Post* contains over twice as much copy as *Liberty*, the factor of boredom through repeated questioning would operate to lower the score for the larger magazine as compared with the smaller.

The writer attempted to use a panel of respondents in a study of audience acceptance of motion pictures. The members of the panel were shown a roster listing the stars, title, and story idea of a number of motion pictures which had not then been released, and were asked to check according to a standardized list of alternatives the likelihood of their seeing each picture. In a follow-up study six months later after all of the pictures had been shown, the members of the panel were given a list on which they checked the titles of those they had actually seen. In connection with this study an attempt was made to find out

how people learn of pictures, the relative importance of word of mouth *vs.* publicity and advertising, etc. To the question: *"How did you first learn of this picture?"* nearly 7 per cent replied with some equivalent of: "Oh, it was on the list you showed me before. Can't you remember?" Obviously being told on the occasion of the first visit that "A large studio in Hollywood would like to know what you think about different motion picture stars and various kinds of stories" made a strong impression on the memory of those panel members. The mere fact that they had been interviewed six months before made them an atypical group on whom to study publicity penetration in relation to attendance.

The motives of the respondent must be considered in appraising the possible effect of repeated interviewing on the honesty and accuracy of the answers given. Steinmetz (8) administered the Strong Vocational Interest Blank under standard conditions. Later he asked the same subjects to fill out the blank in such a way as to qualify them for an arbitrarily chosen occupation. The answers to individual items were greatly changed. The college students could fake their scores sufficiently to make themselves appear qualified for any profession.

Kelly, Miles, and Terman (4) conducted a similar experiment using the M-F test. After the standard administration of the test the subjects were instructed to make themselves more masculine or feminine than they really were. Here too the ability to make oneself appear to be something else was clearly demonstrated. For example, men could make themselves seem more feminine than the average college girl actually is. F. Ruch (7) has shown by a similar method that the Bernreuter Personality Inventory can be influenced in a desired direction. The median extroversion score of a typical group of college students asked to fill in the blank the way they thought a good salesman would, rose to the 98th percentile for extroversion. These three studies leave no room to doubt that the respondent can give a false but consistently patterned report of his impressions if he has any incentive to do so.

Objective facts are much less likely to be influenced by repeated interviewing than are subjective impressions. Especially, impressions that are not "honorable." As the same respondents are contacted time after time they tend to change in two ways. Those who are hostile refuse to co-operate further, while those who are co-operative become more and more at ease, and because of this good *rapport* give information which they have previously withheld.

The repeated interviewing of the same group (panel method) despite its limitations has definite advantages which in certain situations more than offset its disadvantages (1).

In the first place, it is easier to get one good representative sample than to get a number of them. Once the original contact has been made, the follow-up can frequently be done by mail to reduce expense and to save time. Great care, of course, must be taken to see that the panel remains intact. Secondly, it is much easier to relate *change* in behavior to the circumstances of the respondent's life if the panel is used. Fleiss (2), working under the direction of Stanton, reports the results of a panel study using the pantry inventory technique. When a change in buying was noted, the housewife was questioned to determine the cause of the change. Thus changes in buying habits could be related to

radio listening or magazine reading.

In cooperation with Columbia University's Office of Radio Research, Roper (9) used the panel technique in studying changes in political opinion and attitude. A cross-section of the inhabitants of Erie county, Ohio, was followed by means of monthly surveys from May to November. Such a procedure permits change in attitude to be related to its various determining factors. Comparison with the results of standard techniques of interviewing different but equivalent samples indicates that no distortion of results occurred. Apparently the effect of repeated interviewing in the case of political opinions and attitudes is slight.

The evidence on the issues of this paper is too slight to permit any conclusion other than that more research is needed.

REFERENCES

1. Cuthbert, Daniel. *How to Run A Panel.* FREC Pamphlet, Washington, D.C., 1940.
2. Fleiss, Marjorie. The Panel as an Aid in Measuring Effects of Advertising. *Journal of Applied Psychology*, 1940, 24:685-695.
3. Hollingworth, H. L. Judgments of the Comic. *Psychological Review*, 1911, 18: 132-156.
4. Kelly, E. L., Miles, C. C., and Terman, L. M. Ability to Influence One's Score on a Pencil and Paper Test of Personality. *Character and Personality*, 1936, 4: 206-215.
5. Link, H. C. *The New Psychology of Selling and Advertising.* New York: The Macmillan Co., 1932.
6. Maslow, A. H. The Influence of Familiarization on Preference. *Journal of Experimental Psychology*, 1937, 21: 162-180.
7. Ruch, F. L. A Technique for Detecting Attempts to Fafe Performance on Self-Inventory Type of Personality Test. *Studies in Human Personality* (Terman Commemorative Volume). New York: McGraw-Hill Book Co., 1941.
8. Steinmetz, H. L. Measuring Ability to Fake Occupational Interest. *Journal of Applied Psychology*, 1932, 16: 123-130.
9. The People's Choice. *Life*, 1940, 4 (November 11): 95-103.

THE MUSEUM TECHNIQUE APPLIED TO MARKET RESEARCH

By GEORGE K. BENNETT, PH.D.

THE PSYCHOLOGICAL CORPORATION, NEW YORK CITY

THE design and arrangement of commercial exhibits, so far as can be ascertained, is usually planned by individuals who, because of their qualifications in architecture, aesthetics, mechanics, or advertising, have come to be regarded as display experts. Such planning ordinarily produces displays conforming to artistic requirements of arrangement and color. Many are superior examples of commercial art and mechanical ingenuity. The behavior of the visitor in respect to the individual exhibits or the total display situation has received far less attention. Commercial exhibits satisfying those responsible for their design and construction are assumed to be efficient in attracting the attention of the visitor and in conveying to him the intended message.

The commercial display has many points of similarity with the museum. Especially in museums of natural science and industrial arts, much of the content is practically identical with that found in typical advertising exhibits.

Observation of the behavior of museum visitors apparently began with the work of Edward S. Robinson in 1925. In cooperation with the American Association of Museums, he undertook exploratory studies in several museums of art reported in 1928. It was found feasible to follow individual visitors in an inconspicuous manner and to record the total time spent in the museum, the areas visited, and the length of time spent in looking at particular objects. Although the number of visitors observed was small, certain general tendencies characteristic of museum visitors were remarked.

1. During the course of a visit and after a brief "warming up" period the person observed displayed a tendency to stop before a progressively smaller percentage of the pictures encountered and to make progressively shorter stops.
2. The more pictures simultaneously displayed, the smaller the average time spent in looking at each picture.
3. In large museums the likelihood that a visitor will observe any given picture is less than in smaller museums.

Subsequent experimentation by Robinson, Melton and others extended this method of observation (2, 3, 4, 5, 6, 7, 8, 9, 10, 11, 12, 13). Weekday and Sunday visitors have been compared. The effect of position within a museum room has been studied. Accompanied visitors have been compared with unaccompanied persons. The number of stops, the total time spent in the building, and the time spent per stop have been investigated for a variety of situations.

Improved methodology was developed to permit more detailed recording of visitor behavior. Diagrams of the museum were used and the path taken by each visitor was traced by the observer

with symbols to indicate the type of behavior taking place at the stops. A specially constructed portable polygraph took the place of the stopwatches previously used and permitted the observer to obtain detailed records of individual object-times while following his subject for a protracted period.

Space does not permit summarization of the results of the many experiments conducted in this program. However, there have emerged some general principles of museum behavior which operate almost independently of the intrinsic attractiveness of paintings or exhibits. These have been remarked upon by Melton and have repeatedly appeared in later studies of commercial exhibits.

1. In a symmetrically arranged exhibit room there is a strong tendency for the visitor to follow the right hand wall, looking at displays to his right.[1]
2. Exits from a museum room attract the visitor and compete with nearby displays so that stops are less frequent and of less duration before pictures or objects in the vicinity of doors leading from the room. If the visitor encounters an exit before the circuit of the room is completed, he is more likely to depart than to continue in the room.
3. Visitors distribute their attention, usually pausing for brief periods to look at individual objects and then skipping several intervening exhibits before stopping again.

[1] A striking example of the discrepancy between a priori theory and actual behavior has been disclosed by these studies. In a museum gallery exhibiting the work of a single school of art, it is usual to present the paintings in chronological sequence. Such a sequence ordinarily begins at the left of the entrance and proceeds in clockwise rotation. Studies of the behavior of visitors indicate that approximately three out of four proceed from right to left, encountering the sequence in reverse.

Several applications of these techniques have been made to commercial exhibits. The results of these studies indicate that typical exhibits can be redesigned for much greater efficiency when visitor behavior has been observed. Since these studies were made for private clients identifying characteristics have been deleted.

In one display studied, nineteen major exhibits were arranged along the north side of an aisle about ten feet wide and thirty feet long. Five minor exhibits were on the opposite side of the aisle near the eastern end. At the east end of the space was an illuminated wall chart showing in succession the products of the manufacturer. This wall chart and nine of the major exhibits could be put into operation by the visitor. Eleven displays were static and three operated continuously. Four passages to other portions of the building existed, two at each end of the aisle.

Preliminary observation indicated that the majority of visitors entered from the southeast opening. The observer was stationed at this entrance and followed adults who crossed this threshold singly or with one companion. The exhibit diagram was used to indicate the route followed and the portable polygraph to record the time spent before each unit. One hundred thirty-three visitors were observed; of these, one hundred six traversed a major portion of the exhibit. The average time within the exhibit area was 284 seconds of which 214 seconds were spent in observing materials. The mean number of displays observed was 10.2. The average duration of a stop before an operating exhibit was twice that of a stop before a static exhibit.

Five of the displays in this exhibit had definite cycles of operation. Of these, one (the applications of product

display) operated continuously independent of visitor control. The remaining four were so arranged that the visitor, by pressing a button, could initiate the cycle which then continued automatically to completion. In the table below are shown the duration of cycle, average duration of stop, and percentage of cycle observed by the average visitor.

SUMMARY OF MECHANICAL DISPLAY CYCLES

Exhibit	Length of cycle in seconds	Av. duration of stop when display was operating (sec.)	% of cycle observed by visitor
Manufacturing Process	28	27.1	96.8
Chemical Process . . .	94	42.4	45.1
Miniature Theater (with sound)	390	64.0	16.4
Illuminated Wall Chart	391	32.0	8.2
Applications of Product	223	17.3	7.8

On the basis of these observations several suggestions were made to the exhibitor. These included: the use of a directional sign to divert into the exhibit the 20 per cent who chose the nearest exit, the shortening of cycle times to correspond more nearly to visitor habits, and the alternation of operating and static exhibits to diminish crowding.

During 1939, several studies were made in buildings housing commercial exhibits at the New York World's Fair. All but one of these were brief surveys that do not warrant detailed report. Two of the previously remarked tendencies were apparent: the "right-walking" phenomenon and the brevity of observation of single exhibits.

The remaining investigation provided an opportunity to observe behavior under altered conditions since the building arrangements were modified after a sample of visitors had been studied.

This exhibit occupied an entire building. It consisted of a large curved corridor with an auditorium at the right. An impressive, well-designed diorama show lasting about fifteen minutes was presented intermittently in the auditorium. No advertising message was carried by this show, but institutional exhibits were displayed along the corridor and product advertising in a smaller room at the corridor's end. The main exit was adjacent to the room housing the product displays. An auxiliary exit was located at about the center of the corridor.

The first study revealed that the attention of visitors was concentrated upon the main feature and that the advertising displays were neglected. The usual route was such that some corridor exhibits were bypassed and approximately half the visitors left by the auxiliary exit.

On the basis of these findings the routing of visitors to the auditorium was controlled so that each passed several of the corridor exhibits. In front of the auxiliary exit a directional sign was placed showing the way to the smaller "product" room. The net effect of these changes, as determined by the study of a second sample of visitors, was to increase the number of people seeing the corridor displays and entering the smaller room. In spite of these minor effects, the diorama continued to absorb so large a proportion of the visitors as to cast doubt upon the advertising value of this costly exhibit.

The information now available upon the behavior of visitors can be of help to the designers of commercial exhibits. The application of these well-substantiated principles will permit the location of important elements in areas where the visitor will look at them, the placing of exits in such positions that room time will not be diminished and should produce a closer relationship between cycle time and the visitor's interest span.

Because of the importance of the intrinsic attraction values of specific displays and of the competition between

displays, no general formula can be established to predict the effectiveness of an exhibit. However, in the case of important displays, experimental advance installations can be made and studied so that a final arrangement will result which conveys the intended message.

REFERENCES[2]

1. Robinson, E. S. *The Behavior of the Museum Visitor.* Publication of the American Association of Museums, New Series, No. 5, 1928.
2. Robinson, E. S. Psychological Problems of the Science Museum. *Museum News,* 1930, 8: 9-11.
3. Robinson, E. S. Psychological Studies of the Public Museum. *School and Society,* 1931, 33: 1-5.
4. Robinson, E. S. Exit the Typical Visitor. *Journal of Adult Education,* 1931, 3: 418-423.
5. Robinson, E. S. Psychology and Public Policy. *School and Society,* 1933, 37: 1-7.
6. Robinson, E. S. The Psychology of Public Education. *American Journal of Public Health,* 1933, 23: 123-128.
7. Robinson, E. S. Experimental Education in the Museum—a Perspective. *Museum News,* 1933, 11: 6-8.
8. Melton, A. W. Studies of Installation at the Pennsylvania Museum of Art. *Museum News,* 1933, 11: 5-8.
9. Goldberg, N. Experiments in Museum Teaching. *Museum News,* 1933, 11: 6-8.
10. Melton, A. W. Some Behavior Characteristics of Museum Visitors. *Psychological Bulletin,* 1933, 30: 720-721.
11. Melton, A. W. *Problems of Installation in Museums of Art.* Publication of the American Association of Museums, New Series, No. 14, 1935.
12. Melton, A. W. Distribution of Attention in Galleries in a Museum of Science and Industry. *Museum News,* 1936, 14: 6-8.
13. Melton, A. W., Feldman, N. G. and Mason, C. W. *Experimental Studies of the Education of Children in a Museum of Science.* Publication of the American Association of Museums, New Series, No. 15, 1936.

[2] Items are arranged chronologically.

THE ROLE OF PSYCHOLOGICAL INTERPRETATION IN MARKET RESEARCH

By ARTHUR W. KORNHAUSER, Ph.D.

UNIVERSITY OF CHICAGO AND PSYCHOLOGICAL CORPORATION

MARKET research is a very practical matter. It seeks to find improvements in products and in advertising and selling procedures. "Improvements" are those changes which will increase profits.[1] A major portion of the research consequently is concerned with consumer reactions since what is profitable turns largely upon what is popular or acceptable to consumers.

Psychologists have been increasingly active in this field. Their contributions have been chiefly specific techniques—questionnaire and interview methods, rules for sampling, laboratory experimentation, measurements of attitudes and behavior. It is generally agreed that these psychological procedures have proved highly valuable. Where appropriately employed they offer substantial aid in obtaining more valid and reliable results relevant to the practical problems than can be secured without them.[2]

But technical aids for collecting data are not enough. The market research investigator is constantly confronted with the problem of *interpreting* his findings, of drawing *conclusions*. Frequently, to be sure, he strives to evade this responsibility; he tries to believe that the observed facts are his only concern. Obviously, however, this is mere "buck passing"; *someone* has to say what the facts mean. But all too often the interpretations remain on an uncritical snap-judgment level even when the fact-gathering and statistical analysis have been impressively refined. The central object of the present paper is to stress the role of psychology in helping to arrive at more adequate and more carefully considered interpretations — and hence more useful ones. The drawing of valid conclusions is, after all, the paramount objective in market studies.

It should be noted at once that marketing investigations vary greatly in the nature and amount of interpretation demanded. Many inquiries are of a simple fact-finding or census-taking character where the interest of the sponsoring organization lies in the direct implications of the data themselves. For example, the problem is to determine how many people in a given district read a particular newspaper; or it is to ascertain whether a new product is to be sold

[1] The term "market research" is used here in its ordinary present-day connotation to refer to commercial studies for business organization; not in the wide and rarely employed sense of "studies concerning the whole, or any of the elements, of the process of distribution as it relates to the marketing of goods or services" (quoted from the American Marketing Society's book, *The Technique of Marketing Research*, Chart A, p. 2; New York: McGraw-Hill, 1937).

[2] A summary view of these technical contributions may be obtained in such books as the following: A. T. Poffenberger, *Psychology in Advertising*. New York: McGraw-Hill Book Co., 1932; H. C. Link, *The New Psychology of Selling and Advertising*. New York: The Macmillan Co., 1932; H. E. Burtt, *Psychology of Advertising*. New York: Houghton-Mifflin Co., 1938; D. B. Lucas and C. E. Benson, *Psychology for Advertisers*. New York: Harper and Bros., 1930; E. K. Strong, Jr., *Psychological Aspects of Business*. New York: McGraw-Hill Book Co., 1938.

in a green or a yellow wrapper; or the question is simply what brands of cigarettes certain groups of men are buying. At the other extreme studies are conducted which aim to answer such complicated questions as why one automobile is not selling as well as others in its price class and what advertising appeals can best be used to change this. Or the assignment may be to discover people's attitudes toward a public utility company with a view to recommending procedures for increasing goodwill. Insofar as there is a place for psychological interpretation it is clearly greater in these larger research projects.

INTERPRETATION DEPENDS ON COLLECTION OF APPROPRIATE FACTS

Far more than appears on the surface, however, the simpler problems likewise afford opportunity and need for interpretative perspective. This is true in the first place because these "simple" problems are really fragments of more complex ones. Valid conclusions concerning the specific question usually must take into account a wide range of surrounding factors which give the problem its setting. Too often the fragmentary problem assigned the market investigator is not the problem which most needs tackling. Thus the question of detail concerning number of newspaper readers is part of a general problem such as finding the best advertising medium; the users of a certain brand of cigarette are to be counted as a means of measuring the effect of an advertising campaign. Clearly the data collected in the specific study take on new signifiance if they are weighed in terms of their bearing on the broader problem. Likewise the need for supplementary and related facts is then apparent.

But in addition to the necessity for seeing the problem as a whole, it is equally important that the evidence upon each restricted "simple" question be critically interpreted as it pertains to that question. Consider, for example, the simple determination of public preference for the yellow or green package. Suppose a careful sampling of the appropriate population shows 75 per cent preferring the green wrapper. Do the facts then speak for themselves? Is the interpretation automatically that this is the better color to use? Not unless one can plug some loopholes. Perhaps the preference is ephemeral, affected by some temporary fad. Perhaps it is a matter of similarity to a competing brand already familiar and favored. Perhaps the people who prefer the green package also like the yellow reasonably well, whereas most of those choosing the yellow feel intensely negative toward the green. Considerations like these can easily reverse the practical conclusions drawn from simple "cold facts." The broader and more intricate the problem the more urgent does this critical interpretation become.

The central thought in the foregoing is that no amount of expertness in techniques of sampling, collecting data, and statistical analysis can compensate for neglect of the final question: What interpretation is justified? What conclusions does the evidence support? But the further point must now be emphasized that this "final question" is also the *initial* question—or should be. It is not enough to attempt interpretation of whatever data happen to have been collected. The decisive consideration is what facts are needed. Interpretation of results begins with the planning stage of the research. No amount of high-powered interpretation can produce useful conclusions if the wrong data or inadequate data have been collected. In all market research, from the simplest

to the most complex, every step must be taken with the task of interpretation in mind—that is, with a view to arriving at practical, meaningful conclusions.

INTERPRETATION RELATED TO MOTIVES OF BEHAVIOR

More definitely, now, what is the nature of the interpretations in market research? How may they be improved? Above all, what can psychology contribute?

The interpretation of market research findings is the attempt to understand or explain the buying behavior in question and the related factors (both causes and effects) in order to provide the basis for predicting and influencing such behavior. The ultimate purpose is the guiding of marketing policy, the making of practical decisions. However, the immediate problem of interpretation is not "What should we do?" but "What does it mean?"[3] "What does it mean" in the sense of what is involved in this behavior, how does it come about, wherein is it "natural" or the thing to be expected, how may one construe it in order to feel at home with it, to remove all mystery from it? Insofar as one gains this understanding he sees the conduct as similar to more familiar activities, he places it in a causal nexus. As a consequence he feels able to handle practical problems pertaining to this activity and to recommend changes calculated to produce desired results.

Information is accumulated, for example, concerning people's use of chewing gum. They were asked about the frequency with which gum is chewed, about the times and places when they use gum and when they do not, about their likes and dislikes concerning gum-chewing and their reasons for chewing or not chewing, about certain other habits that may be related to this one (smoking, eating candy, etc.), various facts about themselves (age, education, etc.), and so on. Some of the evidence thus collected lends itself to simple interpretation. For example, many persons report that they avoid chewing gum in public, when out in company, in business situations, etc. They also frequently mention that it is bad mannered, does not look nice. These responses are especially numerous among the upper income groups, the better educated, and the middle and older age groups. The clear interpretation of many such lines of evidence is that standards of good taste and fear of social disapproval are especially important deterrents with respect to gum-chewing. Other portions of the data are somewhat less obvious in their interpretation—and some remain quite obscure until further evidence is available. For example, the facts reported about special times when gum is chewed, how long it is chewed, the reasons assigned for chewing, the kind of people reporting these facts and their related habits, support plausible interpretations to the effect that gum-chewing is motivated by direct sensory satisfactions akin to those of eating (oral eroticism?), by its functioning as an emotional outlet and relief of tensions, and by its more deliberate use as a substitute for smoking and candy-eating or in the belief that it aids digestion, purifies the breath, keeps teeth and gums in good condition, and has other similar beneficial effects as advertised. A specific illustration of data the interpretation of which re-

[3] This view is in definite contrast to the treatment by L. O. Brown in his *Market Research and Analysis*. New York: The Ronald Press Co., 1937, Chapter 14. Interpretation there refers to practical recommendations and thus, in the opinion of the present writer, leads to confusion and consequent neglect of understanding and insight (i.e., interpretation as employed here) as a *basis* for the practical recommendations.

mains very uncertain is this: Among men it was found that non-smokers chew more gum than smokers, whereas among women the smokers chew gum more frequently than the non-smokers. Statements of reasons for gum-chewing give some indication that the explanation lies in women's greater tendency to use gum to counteract tobacco taste and "tobacco breath." But further evidence is required on this point.

In these illustrations it is apparent that the interpretations have direct and important implications for advertising and sales. These need not be elaborated here. The practical applications, as stated above, flow from the interpretations but are by no means identical with them. Nor does the success of a given application attest to the validity of the underlying interpretation. Recommended solutions often "work" even where they are based upon unsound interpretations and they often fail, due to faults of application, even where the interpretation is valid. In this paper we shall comment further on the interpretation process itself and shall leave to one side the problem of utilizing the interpretations as guides to action.

Two types of psychological understanding or interpretation are to be distinguished: "personal" and "impersonal." "Personal" understanding consists of putting ourselves in the other fellow's shoes; we interpret his actions on the basis of our own mental states under similar circumstances. The insight is sympathetic, intimate, "intuitive." In the chewing-gum example, we immediately share the motivations which impel people not to chew on certain occasions; we feel the "naturalness" of the behavior, hence "understand" it and can in some degree predict and influence it on the basis of this intimate understanding. By contrast, our "impersonal" understanding is objective and factual, the same sort of causal explanation we employ in the psysical sciences. When we find statistical associations between gum-chewing and age, schooling and occupation, for example, those objective relationships provide added understanding in terms of which we may better predict and control the behavior in question. It is not necessary that the objective data provide the inner "feel" of *why* the relationships occur, though very often they do thus contribute with consequent gain in our ability to deal with the conduct. Clearly the two kinds of interpretation are intertwined in all comprehensive analyses of actual behavior—and both are of utmost importance.

In either case, whether the understanding remains objective and impersonal or proceeds to the inner personal insight (which is perhaps most distinctively called interpretation), factual knowledge is indispensable. Appropriate data about the conduct, the persons, the situations, must be collected and analyzed—in the first instance to establish objective correlations as a basis for prediction and control and in the second instance to provide the descriptive information about the other fellow's adjustments so that we may project ourselves into his position. It is precisely here that the chief need for theoretical guidance, that is psychological guidance, appears. What *kind* of knowledge is to be sought? What data are "appropriate" data? And how do the data fit together into a meaningful picture of the activity under investigation?

In large measure the answers to these questions are provided by ordinary practical judgments based upon experience and upon preliminary inquiries into the particular problem at hand. But that is not to say that more systematic psycho-

logical guidance is not of additional value. A false contrast of facts versus theory often is made the grounds for an opposite conclusion. Emphasis on the need for specific explanatory facts leads some to renounce the use of broad guiding hypotheses, as though the two were antithetical instead of complementary. A number of penetrating market research psychologists, for example, have been so concerned with negative warnings against the search for *basic* motives and *general* dynamisms as explanations of particular market behavior that they have failed altogether to point to the positive usefulness of these general formulations as working hypotheses.[4] Their emphasis, I believe, is in the wrong place. Geunine progress in market research interpretations must come through more, rather than less, willingness to use "theoretical" knowledge of behavior—but always as a source of *suggested* interpretations to be checked upon through inquiry into the facts of the particular case.

This is far from being an "academic" question. Here as elsewhere sound theory promises to be the best guide to practice. What is called for are the best hypotheses that can be formulated concerning the ways in which specific actions are determined. Emphasis must rest upon the motivation or dynamics of the consumer's behavior, but with ample provision for non-motivational determinants of conduct as well. In an earlier article by Dr. Lazarsfeld and the writer[5] an effort was made to suggest the broad lines which such an analysis of action should follow and to point out its usefulness in market research. A brief résumé will serve our present purpose and will permit the singling out of one or two points for special comment.

Each momentary act is analyzed as the joint product of factors in the person and factors in the situation in which he finds himself. The intra-individual factors in turn are divided into "motives" or inner tensions and "mechanisms" or abilities (both motor and intellectual). The mechanisms are not independent of the motives but are directed by them. The analysis aims to inquire into all those motives and mechanisms which are significant in the present behavior and likewise into the related factors in the external situation. But this present behavior is the end-product of a long-continued process. Back of the final response lies the *history* of the motives and mechanisms which now determine conduct. Useful analysis is as vitally concerned with the temporal course of the activity as it is with the single cross-section. The market behavior under investigation must be seen as a sequence of stages along an extended time-line. Each stage is subject to analysis paralleling that applied to the final response. The total analysis seeks to picture and to understand the changing individual as he comes to believe or act in the specific manner which is of present interest to the sponsor of the investigation.

The specific interpretative concepts and working hypotheses dealing with the dynamics of behavior are derived from the whole body of psychology, sometimes from most unexpected places. For example, a principle in the field of

[4] H. C. Link, *op. cit.*, Chapter 4; J. G. Jenkins, *Psychology in Business and Industry*. New York: J. Wiley and Sons, 1935, Chapter 14; D. McGregor, "Motives" as a Tool of Market Research, *Harvard Business Review*, 1940, 19 (1): 42-51; and certain statements in P. F. Lazarsfeld, "The Psychological Aspect of Market Research," *Harvard Business Review*, 1934, 13 (October): 56 ff.

[5] A. W. Kornhauser and P. F. Lazarsfeld, *The Techniques of Market Research from the Standpoint of a Psychologist*, American Management Association, Institute of Management Series, No. 16, 1935.

time perception suggests a type of evidence from which to infer the degree of people's interest or boredom in radio commercials.[6] The principle is simply that the more engrossing the contents of a time interval the shorter it is judged to be. Predominantly, however, the interpretative generalizations are parts of a motivational psychology developed by social and applied psychologists and by psychoanalysts. The key concepts deal with basic motives and their classification, their gratification and frustration; with attitudes and sentiments, including prominently attitudes of self-regard; with Freudian "dynamisms" like repression, displacement, rationalization, projection, identification; with such varied additional matters as habit-formation, deliberation and choice, personality "types" and traits, forms of social interaction, and so on. These concepts may or may not be integrated into a unified psychological theory of action. To the extent that this systematization is feasible, it should contribute significantly to more rounded interpretations which place piecemeal reactions in their broader setting.

The role of these interpretative guides may well be illustrated in the use of motive classifications. It happens that especially sharp attacks have been directed against the use of these motive lists. The criticisms, however, miss the point. They assume that the basic motives must be employed as explanations in lieu of particular and concrete accounts of behavior. Actually, the motives serve rather to suggest what facts to seek and how to interpret the facts. Any reasonably carefully drawn list of general motives calls attention to a range of plausible explanations to be considered. Why does a man purchase the clothes he does? Is it enough to learn that he looks for a suit that fits well or that has style and quality? Certainly these attitudes tell something beyond even more specific explanations pertaining to store location, brand name, window display, etc., though these, too, are significant. But "fit" may, in terms of more general desires, mean primarily bodily comfort to one man, social distinction to another, security from disapprobation to a third, or the increased affection of his lady-love, or a feeling of effectiveness and proper workmanship. The advertiser and salesman need to know the relative importance of such feelings. Appropriate questioning and appropriate interpretation of response guided by basic motive hypotheses can go a long way toward offering answers. Whether the product be suits or life insurance or a presidential candidate, the search for appeals is unlikely to cover the ground unless some comprehensive "check list" of motives is utilized.

For many special problems, more subtle motivational hypotheses must likewise be entertained. Desires only vaguely recognized or altogether unadmitted may at times be inferred from revealing words and actions and may provide invaluable insights—in marketing practice as in more personal affairs. Consider, for example, the elements which make for popularity in numerous movies, comic strips, magazine stories and even parlor games in the light of psychoanalytic emphases upon vicarious and indirect expressions of repressed sexual wishes, guilt-feelings, and aggresive impulses. Or think of the problems of deeply hidden and self-deceptive motivations operating in the patent medicine market. However, the case is almost equally clear in quite ordinary market behavior—in the extravagance and penuriousness manifested in differ-

[6] The method was used by Dr. Paul F. Lazarsfeld in one of his radio inquiries.

ent purchases (or in relation to self and spouse), in disguised efforts to strut without appearing to do so, in over-assertions of independence and in childish acceptance of suggestions. The trouble with psychoanalytic interpretations in the practical world has surely not been a lack of rich and plausible hypotheses but rather the failure to sift evidence to check the validity of the interpretation in each particular application.

In conclusion, a final word is needed to re-emphasize the importance of rounded, over-all interpretations. Any particular explanation, psychoanalytic or otherwise, must be weighed against an array of alternative and supplementary views. Initial interpretations are constantly to be revised and refined in the light of accumulated evidence as the investigation proceeds. A systematic theory of action suggests the comprehensive "check list" of explanations to be examined. Even when most of these suggested explanations are "obvious" they serve most useful purposes. Everyone knows how appallingly easy it is, in the absence of such guidance, to accept certain explanations which occur to us and to overlook others that are equally valid. The investigator requires a psychological map of the whole behavior area involved in order to understand the significance of the pebbles and the nuggets turned up by his spade in any one particular spot.

News Notes

The death of Dr. F. Kuhlmann occurred April 19, 1941, in Saint Paul, Minnesota, after an illness of about two months.

The Kansas Association of Consulting Psychologists held its annual meeting in conjunction with the annual meeting of the Kansas Academy of Science at Manhattan, Kansas, April 4, 1941. Dr. Richard Sears of Leavenworth was elected to membership. The Kansas Association also maintained a booth of psychological exhibits at the Kansas Conference of Social Work in Topeka, April 21-25.

A directory of psychological and psychiatric facilities in Kansas may be obtained upon request to the Kansas Association of Consulting Psychologists.

A paper entitled, "The Genetics of Intelligence," by Professor E. L. Newberry of the University of Kentucky, aroused considerable discussion at a meeting of the Division of Philosophy and Psychology of the Kentucky Academy of Science, which was held April 25 at Eastern State Teachers College, Richmond, Kentucky. Professor Newberry's paper was selected by a committee as the nomination of this section for the King Prize of fifty dollars for the best paper presented at the Academy of Science program each year.

The Kentucky Psychological Association held its annual meeting in conjunction with the Kentucky Education Association in Louisville, April 18. Miss Gertrude Moderow discussed "A Study of the Reading Vocabulary of Slow Seventh Grade Pupils"; Dr. Milton Jensen read a paper, "Psychological Needs of the Teacher of Reading"; and Miss Nancy T. Collins discussed, "The Selection and Use of Standardized Tests in the Classroom." There was a round table discussion of the topic, "What Criteria Should Determine the Selection, Administration, and Interpretation of Psychological Tests Used in the School?"

Throughout the period of the Kentucky Education Association meeting an exhibit of recent books and a file of specimen tests was maintained by the Kentucky Psychological Association, and someone was in attendance for consultation on problems of testing in the schools.

The following officers were elected: Chairman, Dr. Milton B. Jensen; Secretary, Miss Nancy T. Collins.

The work of psychologists in government service during the national emergency was the subject of discussion at a dinner meeting of the Metropolitan New York Association for Applied Psychology held May 2, 1941 at the West Side Y.M.C.A. in New York City. Dr. Alice I. Bryan presented a brief report on her activities with the use of films in discussing problems of national morale with adult groups in local public libraries. Dr. Ladislas Farago, Hungarian journalist and expert on German military affairs, outlined developments in military psychology in Germany since World War I, giving an interesting ac-

count of the systematic use which the Nazis have made of their revisions of earlier American psychological contributions in their attack on the propaganda front in various countries. Dr. Floyd L. Ruch, treasurer of the Committee for National Morale, indicated the effective use which could be made of applied psychology in defending this country against totalitarian propaganda. The Committee for National Morale is made up of publicists, military experts, social scientists, psychologists, industrial workers and others who have come together to study these problems in American national life and to make practical recommendations to the government for taking advantage of the scientific resources of this country.

The Ohio Association for Applied Psychology held its annual meeting in the Social Administration Building of Ohio State University, Saturday, May 10. Officers elected for the coming year are: President, Dr. Oliver A. Ohmann; Secretary-Treasurer, Miss Myra W. Kuenzel; and Representative to the Board of Affiliates of the A.A.A.P., Member-at-Large, Dr. Carl R. Rogers. Following the business meeting there was an informal discussion of problems of personnel selection in public employment offices, problems arising from expansion of industry and problems of psychology and national defense.

The annual meeting of the Pennsylvania Association of Clinical Psychologists was held in Harrisburg, April 28, 1941, with an attendance of sixty-four persons. The morning session was devoted to a discussion of the integration of psychology, psychiatry, and social work, with emphasis on training and competency of workers in these fields as the basis of integration and collaboration. At the business meeting the following officers were elected: President, R. A. Brotemarkle; Vice-President, Marion Monroe; Secretary, William R. Grove; Treasurer, Mary Vanuxem, Members-at-Large of the Executive Committee (terms to expire 1943): G. I. Giardini, Mabel R. Farson, Carroll A. Whitmer. Dr. Brotemarkle led a discussion of psychology and the national emergency at the afternoon session.

The University of Tennessee Psychological Clinic in cooperation with representatives from over thirty public and private welfare agencies, educational institutions, and other organizations formed the Knoxville Psychological Council, Inc., on February 6, 1941. The stated purposes of the Council are: (1) to promote the development of psychological service to the public; (2) to consider ways and means of promoting cooperation and coordination of efforts among agencies already furnishing psychological services to the public; (3) to study the desirability and feasibility of additional psychological services and to consider ways and means of developing such services; and (4) to make the results of its studies available to its supporting agencies and to the public. The Council itself will render no direct psychological service to individuals.

Any individual, agency, group, or institution which is actively engaged in the practice of applied psychology and desires the encouragement and approval of the Council may apply for affiliation. Upon receipt of such application the Council proposes to study the aims and purposes, the qualifications of the staff members, and other facts about the organization. If these meet satisfactory professional standards, the organization will be made an affiliated member of the Council. Periodically reports of the ac-

tivities of all affiliated agencies will be made public.

It is hoped that the Council will serve to promote and coordinate the development of psychological services and that eventually it may serve as an accrediting or standardizing body. One immediate activity will be the establishment of a number of fellowships through which individuals may secure field training with various community agencies under the direction of the University of Tennessee Psychological Clinic.

The Washington-Baltimore Branch of the American Psychological Association held a meeting at the Catholic University of America, Washington, D. C., March 27, 1941, at which the following series of papers by members of the staff of the U. S. Public Health Service was presented: "Psychology and Health Education," by Dr. Arthur Weismann; "Psychological Evaluation of Health Exhibits," by Dr. Mayhew Derryberry; and "The Use of Psychomotor Tests in Industrial Hygiene Research," by Dr. Edmund C. Hammond.

Dr. Jerry W. Carter, Jr. assumed the directorship of the Wichita Child Guidance Center on April 1st, following the retirement of Dr. Edwina A. Cowan who held the directorship of this clinic for the past ten years.

Dr. Richard W. Husband, assistant professor of psychology at the University of Wisconsin, is visiting assistant professor of psychology in the extension service of The Pennsylvania State College. Dr. Husband will be associated with the foreman training courses of the Engineering Defense Program.

PROGRAM OF THE FIFTH ANNUAL PROFESSIONAL CONFERENCE AND BUSINESS MEETING OF THE AMERICAN ASSOCIATION FOR APPLIED PSYCHOLOGY, INC.

NORTHWESTERN UNIVERSITY
EVANSTON, ILLINOIS

Monday, Tuesday, Wednesday and Thursday
September 1, 2, 3, 4, 1941

OFFICERS OF THE ASSOCIATION

President
EDGAR A. DOLL, Vineland Training School

Executive Secretary
C. M. LOUTTIT, Indiana University

Treasurer
JACK W. DUNLAP, University of Rochester

Board of Governors

PRESIDENT, EXECUTIVE SECRETARY, TREASURER, (*ex officio*)

PAUL S. ACHILLES	J. Q. HOLSOPPLE
ALICE I. BRYAN	ELAINE F. KINDER
ALVIN C. EURICH	FRED KUHLMANN
CHARLES A. FORD	MILLICENT POND
DOUGLAS FRYER	C. L. SHARTLE
ARTHUR I. GATES	MORRIS S. VITELES

Program Committee

ALICE I. BRYAN, *Chairman*, ELAINE F. KINDER,
KINSLEY R. SMITH, WM. CLARK TROW
AND THE EXECUTIVE SECRETARY

BRING THIS PROGRAM TO THE MEETING

Program Summary

GENERAL SESSIONS

Monday

10:00-12:00 A.M. Meetings of Professional Committees.
8:30 P.M. Informal Smoker.

Tuesday

9:00-12:00 A.M. Business Meeting, Session I, Committee Reports.
8:00-10:00 P.M. Business Meeting, Session II.

Wednesday

7:00- 9:00 P.M. Annual Dinner and Presidential Address.

CLINICAL SECTION

Monday

3:00- 4:30 P.M. Business Meeting.

Tuesday

9:30-11:30 A.M. Clinical Presentation of Case Material.
2:30- 4:00 P.M. Round Table. Psychological Problems of Later Maturity.

CONSULTING SECTION

Monday

5:00- 6:00 P.M. Business Meeting.

Tuesday

1:30- 3:00 P.M. Round Table. Improving Our Public Relations.
4:00- 6:00 P.M. General Session. Civilian Morale.

EDUCATIONAL SECTION

Monday

12:00- 1:30 P.M. Luncheon and Round Table. The Student Personnel Program in Teacher Training Institutions.
1:30- 3:00 P.M. Business Meeting.
3:30- 6:00 P.M. Excursion to Chicago Bureau of Child Study.

Tuesday

10:30-12:00 A.M. Symposium. The Field of Child Development.
12:15- 2:00 P.M. Luncheon at the CRADLE.
2:30- 4:00 P.M. Panel Discussion. Psychological Workers in the School.

INDUSTRIAL SECTION

Monday

1:30- 3:00 P.M. Business Meeting.
4:00- 5:30 P.M. General Session. Applied Psychology in National Defense.

Tuesday

1:00- 3:00 P.M. Discussion Groups.

MONDAY, SEPTEMBER 1

9:00-10:00 A.M. REGISTRATION. Headquarters, Scott Hall.

10:00-12:00 A.M. MEETINGS OF PROFESSIONAL COMMITTEES.

1. Clinic Certification. ANTHONY J. MITRANO, *Chairman.* Scott Hall, Room 215.
2. Instruction in Applied Psychology. E. R. HENRY, *Chairman.* Harris Hall, Room 104.
3. Professional Employment. JAY L. OTIS, *Chairman.* Harris Hall, Room 306.
4. Professional Ethics. ROBERT T. ROCK, JR., *Chairman.* Harris Hall, Room 308.
5. Public Relations. MARJORIE VAN DE WATER, *Chairman.* University Hall, Room 102.
6. Relations with the Educational Profession. M. R. TRABUE, *Chairman.* University Hall, Room 106.
7. Relations with the Library Profession. ALICE I. BRYAN, *Chairman.* University Hall, Room 201.
8. Relations with the Medical Profession. LOWELL S. SELLING, *Chairman.* University Hall, Room 202.
9. Relations with the Social Work Profession. EDWINA A. COWAN, *Chairman.* University Hall, Room 203.
10. Standardized Tests. HENRY E. GARRETT, *Chairman.* Harris Hall, Room 306.
11. Technical Manuals. HERBERT A. TOOPS, *Chairman.* Harris Hall, Room 313.

(*Note*: The meetings of these committees will be open to any member who is interested in the Committe's function.)

10:00-12:00 A.M. MEETINGS OF SECTION COUNCILS.

10:00-12:00 A.M. BOARD OF EDITORS. Scott Hall, Room 216.

10:00-12:00 A.M. BOARD OF AFFILIATES. Scott Hall, Room 217.

12:00- 1:30 P.M. LUNCHEON AND ROUND TABLE. THE STUDENT PERSONNEL PROGRAM IN TEACHER TRAINING INSTITUTIONS. EDMUND G. WILLIAMSON, *Chairman.*
Vera Megowen's, 1641 Orrington Avenue. (Make reservations by mail with DR. W. C. TROW, University of Michigan, or at registration desk upon arrival.)
Participants: RALPH C. BEDELL, L. L. LOVE, DEWEY STUIT.

1:30- 3:00 P.M. BUSINESS MEETING, EDUCATIONAL SECTION. ARTHUR I. GATES, *Chairman.* University Hall, Room 104.

1:30- 3:00 P.M. BUSINESS MEETING, INDUSTRIAL SECTION. ARTHUR W. KORNHAUSER, *Chairman.* Fisk Hall, Room 101.

3:00- 4:30 P.M. BUSINESS MEETING, CLINICAL SECTION. ELAINE F. KINDER, *Chairman.* Harris Hall, Room 107.

3:30- 6:00 P.M. EXCURSION, EDUCATIONAL SECTION. To the CHICAGO BUREAU OF CHILD STUDY. (Address, 228 North La Salle Street.) GRACE MUNSON, *Hostess.*

Program: Tea, Exhibits, Motion Pictures.

(This excursion will start from Scott Hall. Make reservations at registration desk upon arrival.)

4:00- 5:30 P.M. GENERAL SESSION. APPLIED PSYCHOLOGY IN NATIONAL DEFENSE. Fisk Auditorium.

Participants in this symposium will be announced at the meeting. It is planned to present brief reports on applications of psychology in defense activities.

(Session organized by Industrial Section.)

5:00- 6:00 P.M. BUSINESS MEETING, CONSULTING SECTION. PAUL S. ACHILLES, *Chairman.* Harris Hall, Room 202.

8:30 P.M. INFORMAL SMOKER. (Bridge and dancing). Scott Hall.

TUESDAY, SEPTEMBER 2

9:00-12:00 A.M. ASSOCIATION BUSINESS MEETING, SESSION I. REPORTS OF STANDING COMMITTEES. EDGAR A. DOLL, *Chairman.* Swift Engineering Building, Room 111.

This session constitutes a part of the Annual Business Meeting. It is hoped there will be a good attendance with active discussion. A timetable of reports will be posted on the Bulletin Board in Scott Hall.

9:30-11:30 A.M. CLINICAL SESSION. JOHN J. B. MORGAN, *Chairman.* Harris Hall, Room 202.

Participants:

Presentation of clinical case material: AUGUSTA T. JAMESON, JERRY W. CARTER, JR.

Discussants: CATHERINE COX MILES, GILBERT J. RICH

10:30-12:00 A.M. SYMPOSIUM. THE FIELD OF CHILD DEVELOPMENT. HAROLD E. JONES, *Chairman.* Harris Hall, Room 107.

Participants:

HAROLD E. JONES, The Problems of Child Development.

JOHN E. ANDERSON, Relationships to Psychology.

HORACE B. ENGLISH, Relationships to Education.

GEORGE D. STODDARD, Needed Integrations in the Field of Child Development.

12:15- 2:00 P.M. LUNCHEON, EDUCATIONAL SECTION. At the CRADLE, 2049 Ridge Avenue (entrance on Simpson Street). MRS. WILLIAM B. WALRATH, Managing Director, *Hostess.* Exhibits and observation.
(Make reservations at registration desk upon arrival.)

1:00- 3:00 P.M. DISCUSSION GROUPS, INDUSTRIAL SECTION.
The Industrial Section is introducing a new form of meeting session consisting of small discussion groups. These will bring together small numbers of individuals mutually interested in discussing various topics in informal fashion. See bulletin board for list of topics and places. Members of the Industrial Section wishing to present a topic for discussion should correspond with Dr. K. R. Smith, Program Chairman, as early as possible.

1:30- 3:00 P.M. ROUND TABLE. IMPROVING OUR PUBLIC RELATIONS. ROBERT A. BROTEMARKLE, *Chairman.* Harris Hall, Room 202.
Participants:
FRANK STANTON, Press and Radio.
HAROLD E. BURTT, Business and Industry.
STEUART H. BRITT, Legislation and Certification.
RENSIS LIKERT, Governmental Agencies.
ROSE G. ANDERSON, Educational and Vocational Guidance.

2:30- 4:00 P.M. PANEL DISCUSSION. PSYCHOLOGICAL WORKERS IN THE SCHOOL. C. GILBERT WRENN, *Chairman.* Scott Hall, Room 215.
Participants: O. L. CRISSEY, F. H. FINCH, LEOTA BOETTICHER, DANIEL E. PRESCOTT.

2:30- 4:00 P.M. ROUND TABLE. PSYCHOLOGICAL PROBLEMS OF LATER MATURITY. GEORGE LAWTON, *Chairman.* Scott Hall, Room 217.
Participants: F. H. FINCH, JEANNE G. GILBERT, A. R. GILLILAND, IVA L. PETERS, MORTON A. SEIDENFELD, LOWELL S. SELLING, PERCIVAL M. SYMONDS.

4:00- 6:00 P.M. GENERAL SESSION. CIVILIAN MORALE. H. MELTZER, *Chairman.* Fisk Auditorium.
Participants:
GORDON W. ALLPORT, The Democratic Basis for Civilian Morale.
ROBERT N. MCMURRY, Industrial Morale.
MARY H. S. HAYES, Youth Problems and Morale.
ALICE I. BRYAN, Adult Education and Morale.

LAURANCE F. SHAFFER, Mental Hygiene Aspects of Morale.
FLOYD L. RUCH, Public Opinion in Civilian Morale.
GOODWIN WATSON, Summary and Discussion.
(Session organized by Consulting Section.)

8:00-10:00 P.M. ASSOCIATION BUSINESS MEETING. SESSION II. EDGAR A. DOLL, *Chairman.* Swift Engineering Building, Room 111.

WEDNESDAY, SEPTEMBER 3

NOTE.—Research programs arranged in cooperation with the American Psychological Association. Starred (*) papers were submitted to the American Association for Applied Psychology.

CLINICAL PSYCHOLOGY

WEDNESDAY, SEPTEMBER 3, 9:00 A.M.

DOUGLAS H. FRYER, *Chairman*

9:00 *The development and evaluation of a measure of counseling interview procedure.* E. H. PORTER, JR., Ohio State University. [15 min.]

9:20 *Constructive therapy for adults.* D. H. YATES, San Jose State College. [15 min.]

9:40 *Aptitude determination in the field of clinical psychology.* M. C. MYERS, University of California. [10 min.]

9:55 *A statistical analysis of the factors in psycho-clinical diagnosis and procedure.* R. M. STOGDILL, Ohio State Bureau of Juvenile Research. [15 min.]

10:15 *A biological and psychological report on a pair of monozygotic Mongoloid twins.* NORMAN FORD, University of Toronto and SYLVIA FRUMKIN, Greystone Park, N.J. [10 min., slides.]

*10:30 *Psycho-somatographs of a case of hypo-gonadism before and after treatment.* H. W. MARTIN, Emory University. [10 min., slides.]

*10:45 *Fetal irradiation and microcephalic idiocy: report of a case in litigation.* EDGAR A. DOLL, Vineland Training School. [15 min., slides.]

*11:05 *A quantitative investigation of the variability of reaction as a measure of emotional stability.* PAUL M. FITTS, University of Tennessee. [15 min., slides.]

11:25 *The quantitative study of language behavior.* WENDELL JOHNSON, University of Iowa. [15 min., slides.]

WEDNESDAY, SEPTEMBER 3

7:00- 8:00 P.M. ANNUAL DINNER OF THE ASSOCIATION. HORACE B. ENGLISH, *Presiding.*

8:00- 9:00 P.M. PRESIDENTIAL ADDRESS. EDGAR A. DOLL. Scientific Freedom.

(Dress optional. Members and guests not attending the dinner are invited to hear the Presidential Address.)

THURSDAY, SEPTEMBER 4

EDUCATIONAL PSYCHOLOGY

THURSDAY, SEPTEMBER 4, 9:00 A.M.

EDGAR A. DOLL, *Chairman*

9:00 *Student evaluation of knowledge of college aptitude test score.* R. K. COMPTON, South Dakota State College. [10 min.]

9:15 *Factors bearing on academic achievement among institutionalized mentally deficient children of moron and borderline level.* RUTH T. MELCHER, Wayne County Training School. [15 min., slides.]

9:35 *Psychopathological causes of inefficiency in our colleges.* JAMES L. GRAHAM, Lehigh University. [15 min.]

9:55 *Primary mental abilities and scholastic success in the divisional studies at the University of Chicago.* K. S. YUM, University of Chicago. [15 min., slides.]

10:15 *Differential prediction of college achievement in the College of Letters and Science at the University of Wisconsin.* C. H. RUEDISILI, University of Wisconsin. [15 min.]

*10:35 *A socialized service-laboratory program in educational psychology.* S. L. PRESSEY, Ohio State University. [15 min.]

10:55 *Some effects of ten years of psychological service in a public school system.* W. L. WILKINS, Shorewood Public Schools. [10 min.]

11:10 *The stability of measured interests of adolescent girls over a two-year period.* LEONA E. TYLER, University of Oregon. [10 min.]

FRIDAY, SEPTEMBER 5

INTELLIGENCE TESTS

FRIDAY, SEPTEMBER 5, 1:15 P.M.

R. S. WOODWORTH, *Chairman*

1:15 *Some theoretical issues in adult intelligence testing.* RAYMOND B. CATTELL, Clark University. [15 min.]

1:35 *A comparison of the complete and abbreviated Stanford-Binet and of clusters of items within the test for adult offenders.* L. I. MAIZLISH, Diagnostic Depot, Joilet, Illinois. [10 min.]

1:50 *Clinical evidence concerning the nature of normal adult test performance.* D. O. HEBB, Quenn's University. [15 min., slides.]

*2:10 *The revised Stanford-Binet Scale viewed as a point scale.* C. H. GROWDON, Ohio State Bureau of Juvenile Research. [15 min.]

*2:30 *Children with inferior social histories: their mental development in foster homes.* H. M. SKEELS, University of Iowa. [15 min.]

2:50 *Seasonal variations in mental growth.* HAROLD E. JONES, University of California. [10 min., slides.]

3:05 *A test for measuring the mentality of infants.* A. R. GILLILAND, and J. J. B. MORGAN, Northwestern University. [15 min., slides.]

3:25 *A comparative study of the intelligence of Jewish and Scandinavian kindergarten children.* FRED BROWN, Minneapolis Public Schools. [15 min.]

AMERICAN ASSOCIATION FOR APPLIED PSYCHOLOGY, INC.

The object of this Association is the promotion and development of the service which psychology as a science can render society in its application in business, education, industry, law, medicine, mental hygiene, social welfare and related fields. To that end it is concerned with the development of psychologists' service and practice at a high level, with the encouragement and facilitation of research, and with the promotion of high professional and ethical standards. Its particular business shall be: the definition of professional and ethical standards, the improvement of training facilities in applied psychology, the conduct of meetings for the interchange of information among its members, the issuance of publications for both its members and the general public, the furtherance of cooperation with non-psychological groups, the development of new professional opportunities, and the transaction of such other business as is consistent with its primary purposes or is authorized by its Articles of Incorporation or its By-Laws.

EXECUTIVE OFFICERS

President, EDGAR A. DOLL, The Training School, Vineland, N.J.
Executive Secretary, C. M. LOURTIT, PH.D., Indiana University, Bloomington, Ind.
Treasurer, J. W. DUNLAP, PH.D., University of Rochester, Rochester, N.Y.

BOARD OF GOVERNORS

The Board of Governors consists of the President, the Executive Secretary, the Treasurer, the Chairman of each Section, the Secretary of each Section, the Chairman of the Board of Affiliates, and the Chairman of the Board of Editors.

SECTION OFFICERS

Section of Clinical Psychology

Vice-President, F. EUHLMANN, PH.D., Division of Social Welfare, Department of Social Security, St. Paul, Minn.
Secretary, ELAINE F. KINDER, PH.D., Letchworth Village, Thiells, N.Y.

Section of Consulting Psychology

Vice-President, P. S. ACHILLES, PH.D., The Psychological Corporation, New York, N.Y.
Secretary, ALICE I. BRYAN, PH.D., School of Library Service, Columbia University, New York, N.Y.

Section of Educational Psychology

Vice-President, ARTHUR I. GATES, PH.D., Teachers College, Columbia University
Secretary, A. C. ELDRICH, PH.D., Stanford University, Calif.

Section of Industrial and Business Psychology

Vice-President, A. W. KORNHAUSER, PH.D., University of Chicago
Board Representative, M. S. VITELES, PH.D.
Secretary, MILLICENT POND, PH.D., Scovill Manufacturing Co., Waterbury, Conn.

OFFICERS OF BOARD OF EDITORS

Chairman, DOUGLAS FRYER, PH.D., New York University, University Heights, New York, N.Y.
Secretary, C. L. SHARTLE, PH.D., Social Security Board, Washington, D.C.

SECTION REPRESENTATIVES TO BOARD OF EDITORS

Section of Clinical Psychology: E. A. DOLL, PH.D., GERTRUDE HILDRETH, PH.D., C. M. LOURTIT, PH.D., BERTHA M. LUCKEY, PH.D.
Section of Consulting Psychology: KARL R. FERNALD, PH.D., DOUGLAS FRYER, PH.D., R. H. PATTERSON, PH.D., F. L. WELLS, PH.D.
Section of Educational Psychology: R. C. BITTNER, PH.D., H. H. REMMERS, PH.D., C. R. ROGERS, PH.D., PERCIVAL M. SYMONDS, PH.D.
Section of Industrial and Business Psychology: H. E. BURTT, PH.D., H. M. JOHNSON, PH.D., C. L. SHARTLE, PH.D., SAMM S. SHELLOW, PH.D.

REGIONAL SOCIETIES AFFILIATED WITH THE AMERICAN ASSOCIATION FOR APPLIED PSYCHOLOGY, INC.

Listed in Order of Affiliation

OFFICERS OF BOARD OF AFFILIATES

Chairman, J. G. Shakespeare, Ph.D., Department of Institutions and Agencies, Trenton, N. J.
Secretary, C. A. Penn, Ph.D., Temple University, Philadelphia, Pa.

SECTION REPRESENTATIVES TO BOARD OF AFFILIATES

N. Yerkes, Ph.D., J. G. Shakespeare, Ph.D., C. A. Penn, Ph.D., Elisabeth E. Lord, Ph.D., Harriet E. O'Shea, Ph.D., W. A. Kelly, Ph.D., G. H. Deabuck, Ph.D., Myra W. Kuehne, B.A., S. A. Reid, Ph.D., Sidney Benjamin Berry, Ph.D. (Listed in the order of the affiliating societies represented.)

NEW JERSEY ASSOCIATION OF PSYCHOLOGISTS, 1937

President, Sidney Sanderson, Ph.D., Rutgers University
Secretary, Mildred E. Germershausen, B.A., North Jersey Training School, Little Falls, N.J.

NEW YORK STATE ASSOCIATION FOR APPLIED PSYCHOLOGY, INC., 1937

President, Ethel L. Cornell, Ph.D., New York State Education Department
Secretary, Dorothea MacCartney, Ph.D., Fordham University, 302 Broadway, New York, N.Y.

PENNSYLVANIA ASSOCIATION OF CLINICAL PSYCHOLOGISTS, 1938

President, Robert A. Brotemarkle, Ph.D., University of Pennsylvania
Secretary, Wilbert E. Grove, Ph.D., Behavior Clinic, Pittsburgh, Pa.

BOSTON SOCIETY OF CLINICAL PSYCHOLOGISTS, 1939

President, Elisabeth E. Lord, Ph.D., Children's Hospital, Brookline, Mass.
Secretary, Stanley G. Estes, Ph.D., Northeastern University, Boston, Mass.

INDIANA ASSOCIATION OF CLINICAL PSYCHOLOGISTS, INC., 1939

President, Harriet E. O'Shea, Ph.D., Purdue University
Secretary, Hazel Hansford Stevens, Ph.D., Family Welfare Society, Indianapolis, Ind.

KANSAS STATE ASSOCIATION OF CONSULTING PSYCHOLOGISTS, 1939

President, Erwina A. Cowan, Ph.D., Wichita Child Guidance Center, Wichita, Kan.
Secretary, Robert L. Erickson, Ph.D., State Department of Social Welfare, Topeka, Kan.

KENTUCKY PSYCHOLOGICAL ASSOCIATION, 1939

Chairman, Milton B. Jensen, Ph.D., Louisville, Ky.
Secretary, Nancy T. Collins, Louisville Public Schools

OHIO ASSOCIATION FOR APPLIED PSYCHOLOGY, 1939

President, George A. Ohmann, Ph.D., Cleveland College
Secretary, Myra W. Kuehne, B.A., 5054 Madison Road, Cincinnati, O.

WISCONSIN ASSOCIATION FOR APPLIED PSYCHOLOGY, 1939

President, Samuel A. Kirk, Ph.D., State Teachers College, Milwaukee, Wis.
Secretary, Ramona Mellenbruch, Ph.D., Milwaukee Public Schools, Milwaukee, Wis.

WASHINGTON-BALTIMORE BRANCH OF AMERICAN PSYCHOLOGICAL ASSOCIATION, 1940

President, Vernon P. Shimer, Ph.D., Waverly Terrace, Baltimore, Md.
Secretary, Rogers M. Bellows, Ph.D., University of Maryland